Introduction to Groups, Invariants and Particles

Frank W. K. Firk, Professor Emeritus of Physics, Yale University

2000

CONTENTS

PREFACE

This introduction to Group Theory, with its emphasis on Lie Groups and their application to the study of symmetries of the fundamental constituents of matter, has its origin in a one-semester course that I taught at Yale University for more than ten years. The course was developed for Seniors, and advanced Juniors, majoring in the Physical Sciences. The students had generally completed the core courses for their majors, and had taken intermediate level courses in Linear Algebra, Real and Complex Analysis, Ordinary Linear Differential Equations, and some of the Special Functions of Physics. Group Theory was not a mathematical requirement for a degree in the Physical Sciences. The majority of existing undergraduate textbooks on Group Theory and its applications in Physics tend to be either highly qualitative or highly mathematical. The purpose of this introduction is to steer a middle course that provides the student with a sound mathematical basis for studying the symmetry properties of the fundamental particles. It is not generally appreciated by Physicists that continuous transformation groups (Lie Groups) originated in the Theory of Differential Equations. The infinitesimal generators of Lie Groups therefore have forms that involve *differential operators* and their *commutators*, and these operators and their algebraic properties have found, and continue to find, a natural place in the development of Quantum Physics.

Guilford, CT.

June, 2000.

1

INTRODUCTION

The notion of geometrical symmetry in Art and in Nature is a familiar one. In Modern Physics, this notion has evolved to include symmetries of an abstract kind. These new symmetries play an essential part in the theories of the microstructure of matter. The basic symmetries found in Nature seem to originate in the mathematical structure of the laws themselves, laws that govern the motions of the galaxies on the one hand and the motions of quarks in nucleons on the other.

In the Newtonian era, the laws of Nature were deduced from a small number of imperfect observations by a small number of renowned scientists and mathematicians. It was not until the Einsteinian era, however, that the significance of the symmetries associated with the laws was fully appreciated. The discovery of space-time symmetries has led to the widely-held belief that the laws of Nature can be derived from symmetry, or invariance, principles. Our incomplete knowledge of the fundamental interactions means that we are not yet in a position to confirm this belief. We therefore use arguments based on empirically established laws and restricted symmetry principles to guide us in our search for the fundamental symmetries. Frequently, it is important to understand why the symmetry of a system is observed to be broken.

In Geometry, an object with a definite shape, size, location, and orientation constitutes a state whose symmetry properties, or invariants,

are to be studied. Any transformation that leaves the state unchanged in form is called a symmetry transformation. The greater the number of symmetry transformations that a state can undergo, the higher its symmetry. If the number of conditions that define the state is reduced then the symmetry of the state is increased. For example, an object characterized by oblateness alone is symmetric under all transformations except a change of shape.

In describing the symmetry of a state of the most general kind (not simply geometric), the algebraic structure of the set of symmetry operators must be given; it is not sufficient to give the number of operations, and nothing else. The *law of combination* of the operators must be stated. It is the *algebraic group* that fully characterizes the symmetry of the general state.

The *Theory of Groups* came about unexpectedly. Galois showed that an equation of degree n, where n is an integer greater than or equal to five cannot, in general, be solved by algebraic means. In the course of this great work, he developed the ideas of Lagrange, Ruffini, and Abel and introduced the concept of a *group*. Galois discussed the functional relationships among the roots of an equation, and showed that the relationships have symmetries associated with them under permutations of the roots.

The operators that transform one functional relationship into another are elements of a set that is characteristic of the equation; the set of operators is called the Galois group of the equation.

In the 1850's, Cayley showed that every finite group is isomorphic to a certain permutation group. The geometrical symmetries of crystals are described in terms of finite groups. These symmetries are discussed in many standard works (see bibliography) and therefore, they will not be discussed in this book.

In the brief period between 1924 and 1928, Quantum Mechanics was developed. Almost immediately, it was recognized by Weyl, and by Wigner, that certain parts of Group Theory could be used as a powerful analytical tool in Quantum Physics. Their ideas have been developed over the decades in many areas that range from the Theory of Solids to Particle Physics.

The essential role played by groups that are characterized by parameters that vary continuously in a given range was first emphasized by Wigner. These groups are known as *Lie Groups.* They have become increasingly important in many branches of contemporary physics, particularly Nuclear and Particle Physics. Fifty years after Galois had introduced the concept of a group in the Theory of Equations, Lie introduced the concept of a continuous transformation group in the Theory of Differential Equations. Lie's theory unified many of the disconnected methods of solving differential equations that had evolved over a period of

two hundred years. Infinitesimal unitary transformations play a key role in discussions of the fundamental conservation laws of Physics.

In Classical Dynamics, the invariance of the equations of motion of a particle, or system of particles, under the Galilean transformation is a basic part of everyday relativity. The search for the transformation that leaves Maxwell's equations of Electromagnetism unchanged in form (invariant) under a linear transformation of the space-time coordinates, led to the discovery of the Lorentz transformation. The fundamental importance of this transformation, and its related invariants, cannot be overstated.

2

GALOIS GROUPS

In the early 19th - century, Abel proved that it is not possible to solve the general polynomial equation of degree greater than four by algebraic means. He attempted to characterize all equations that can be solved by radicals. Abel did not solve this fundamental problem. The problem was taken up and solved by one of the greatest innovators in Mathematics, namely, Galois.

2.1. Solving cubic equations

The main ideas of the Galois procedure in the Theory of Equations, and their relationship to later developments in Mathematics and Physics, can be introduced in a plausible way by considering the standard problem of solving a cubic equation.

Consider solutions of the general cubic equation

$Ax^3 + 3Bx^2 + 3Cx + D = 0$, where $A - D$ are rational constants.

If the substitution $y = Ax + B$ is made, the equation becomes

$$y^3 + 3Hy + G = 0$$

where

$$H = AC - B^2$$

and

$$G = A^2D - 3ABC + 2B^3.$$

The cubic has three real roots if $G^2 + 4H^3 < 0$ and two imaginary roots if $G^2 + 4H^3 > 0$. (See any standard work on the Theory of Equations).

If all the roots are real, a trigonometrical method can be used to obtain the solutions, as follows:

the Fourier series of $\cos^3 u$ is

$$\cos^3 u = (3/4)\cos u + (1/4)\cos 3u.$$

Putting

$$y = s\cos u \text{ in the equation } y^3 + 3Hy + G = 0$$

$$(s > 0),$$

gives

$$\cos^3 u + (3H/s^2)\cos u + G/s^3 = 0.$$

Comparing the Fourier series with this equation leads to

$$s = 2\sqrt{(-H)}$$

and

$$\cos 3u = -4G/s^3.$$

If v is any value of u satisfying $\cos 3u = -4G/s^3$, the general solution is

$$3u = 2n\pi \pm 3v, \text{ where n is an integer.}$$

Three different values of cosu are given by

$$u = v, \text{ and } 2\pi/3 \pm v.$$

The three solutions of the given cubic equation are then

$$s\cos v, \text{ and } s\cos(2\pi/3 \pm v).$$

Consider solutions of the equation

$$x^3 - 3x + 1 = 0.$$

In this case,

$$H = -1 \text{ and } G^2 + 4H^3 = -3.$$

All the roots are therefore real, and they are given by solving

$$\cos 3u = -4G/s^3 = -4(1/8) = -1/2$$

or,

$$3u = \cos^{-1}(-1/2).$$

The values of u are therefore $2\pi/9$, $4\pi/9$, and $8\pi/9$, and the roots are

$$x_1 = 2\cos(2\pi/9),\ x_2 = 2\cos(4\pi/9), \text{ and } x_3 = 2\cos(8\pi/9).$$

2.2. Symmetries of the roots

The roots x_1, x_2, and x_3 exhibit a simple pattern. Relationships among them can be readily found by writing them in the complex form $2\cos\theta = e^{i\theta} + e^{-i\theta}$ where $\theta = 2\pi/9$ so that

$$x_1 = e^{i\theta} + e^{-i\theta},$$

$$x_2 = e^{2i\theta} + e^{-2i\theta},$$

and

$$x_3 = e^{4i\theta} + e^{-4i\theta}.$$

Squaring these values gives

$$x_1^2 = x_2 + 2,$$

$$x_2^2 = x_3 + 2,$$

and

$$x_3^2 = x_1 + 2.$$

The relationships among the roots have the functional form $f(x_1,x_2,x_3) = 0$. Other relationships exist; for example, by considering the sum of the roots we find

$$x_1 + x_2^2 + x_2 - 2 = 0$$

$$x_2 + x_3^2 + x_3 - 2 = 0,$$

and

$$x_3 + x_1^2 + x_1 - 2 = 0.$$

Transformations from one root to another can be made by doubling-the-angle, θ.

The functional relationships among the roots have an algebraic symmetry associated with them under interchanges (substitutions) of the roots. If Ω is the operator that changes $f(x_1,x_2,x_3)$ into $f(x_2,x_3,x_1)$ then

$$\Omega f(x_1,x_2,x_3) \rightarrow f(x_2,x_3,x_1),$$

$$\Omega^2 f(x_1,x_2,x_3) \rightarrow f(x_3,x_1,x_2),$$

and

$$\Omega^3 f(x_1,x_2,x_3) \rightarrow f(x_1,x_2,x_3).$$

The operator $\Omega^3 = \mathbf{I}$, is the identity.

In the present case,

$$\Omega(x_1^2 - x_2 - 2) = (x_2^2 - x_3 - 2) = 0,$$

and

$$\Omega^2(x_1^2 - x_2 - 2) = (x_3^2 - x_1 - 2) = 0.$$

2.3. The Galois group of an equation.

The set of operators $\{\mathbf{I}, \Omega, \Omega^2\}$ introduced above, is called the Galois group of the equation $x^3 - 3x + 1 = 0$. (It will be shown later that it is isomorphic to the cyclic group, C_3).

The elements of a Galois group are operators that interchange the roots of an equation in such a way that the transformed functional relationships are true relationships. For example, if the equation

$$x_1 + x_2^2 + x_2 - 2 = 0$$

is valid, then so is

$$\Omega(x_1 + x_2^2 + x_2 - 2) = x_2 + x_3^2 + x_3 - 2 = 0.$$

True functional relationships are polynomials with rational coefficients.

2.4. Algebraic fields

We now consider the Galois procedure in a more general way. An algebraic solution of the general nth - degree polynomial

$$a_o x^n + a_1 x^{n-1} + \ldots a_n = 0$$

is given in terms of the coefficients a_i using a finite number of operations (+,-,×,÷,√). The term "solution by radicals" is sometimes used because the

operation of extracting a square root is included in the process. If an infinite number of operations is allowed, solutions of the general polynomial can be obtained using transcendental functions. The coefficients a_i necessarily belong to a *field* which is closed under the rational operations. If the field is the set of rational numbers, Q, we need to know whether or not the solutions of a given equation belong to Q. For example, if

$$x^2 - 3 = 0$$

we see that the coefficient -3 belongs to Q, whereas the roots of the equation, $x_i = \pm \sqrt{3}$, do not. It is therefore necessary to *extend* Q to Q', (say) by adjoining numbers of the form $a\sqrt{3}$ to Q, where a is in Q.

In discussing the cubic equation $x^3 - 3x + 1 = 0$ in **2.2**, we found certain functions of the roots $f(x_1,x_2,x_3) = 0$ that are symmetric under permutations of the roots. The symmetry operators formed the Galois group of the equation.

For a general polynomial:

$$x^n + a_1x^{n-1} + a_2x^{n-2} + .. a_n = 0,$$

functional relations of the roots are given in terms of the coefficients in the standard way

$$x_1 + x_2 + x_3 .. \quad .. + x_n = -a_1$$

$$x_1x_2 + x_1x_3 + .. x_2x_3 + x_2x_4 + ..+ x_{n-1}x_n = a_2$$

$$x_1x_2x_3 + x_2x_3x_4 + .. \quad .. + x_{n-2}x_{n-1}x_n = -a_3$$

. .

$x_1x_2x_3$ .. .. $x_{n-1}x_n = \pm a_n$.

Other symmetric functions of the roots can be written in terms of these basic symmetric polynomials and, therefore, in terms of the coefficients. Rational symmetric functions also can be constructed that involve the roots and the coefficients of a given equation. For example, consider the quartic

$$x^4 + a_2x^2 + a_4 = 0.$$

The roots of this equation satisfy the equations

$$x_1 + x_2 + x_3 + x_4 = 0$$

$$x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 = a_2$$

$$x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 = 0$$

$$x_1x_2x_3x_4 = a_4.$$

We can form any rational symmetric expression from these basic equations (for example, $(3a_4^3 - 2a_2)/2a_4^2 = f(x_1,x_2,x_3,x_4)$). In general, every rational symmetric function that belongs to the field F of the coefficients, a_i, of a general polynomial equation can be written rationally in terms of the coefficients.

The Galois group, Ga, of an equation associated with a field F therefore has the property that if a rational function of the roots of the equation is invariant under all permutations of Ga, then it is equal to a quantity in F.

Whether or not an algebraic equation can be broken down into simpler equations is important in the theory of equations. Consider, for example, the equation

$$x^6 = 3.$$

It can be solved by writing $x^3 = y$, $y^2 = 3$ or

$$x = (\sqrt{3})^{1/3}.$$

To solve the equation, it is necessary to calculate square and cube roots ¾ not sixth roots. The equation $x^6 = 3$ is said to be compound (it can be broken down into simpler equations), whereas $x^2 = 3$ is said to be atomic. The atomic properties of the Galois group of an equation reveal the atomic nature of the equation, itself. (In Chapter **5**, it will be seen that a group is atomic ("simple") if it contains no proper invariant subgroups).

The determination of the Galois groups associated with an arbitrary polynomial with unknown roots is far from straightforward. We can gain some insight into the Galois method, however, by studying the group structure of the quartic

$$x^4 + a_2x^2 + a_4 = 0$$

with known roots

$$x_1 = ((-a_2 + \mu)/2)^{1/2}\ ,\ x_2 = -x_1,$$

$$x_3 = ((-a_2 - \mu)/2)^{1/2}\ ,\ x_4 = -x_3,$$

where

$$\mu = (a_2{}^2 - 4a_4)^{1/2}.$$

The field F of the quartic equation contains the rationals Q, and the rational expressions formed from the coefficients a_2 and a_4.

The relations

$$x_1 + x_2 = x_3 + x_4 = 0$$

are in the field F.

Only eight of the 4! possible permutations of the roots leave these relations invariant in F; they are

$$\left\{ P_1 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_1 & x_2 & x_3 & x_4 \end{pmatrix}, \; P_2 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_1 & x_2 & x_4 & x_3 \end{pmatrix}, \; P_3 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_1 & x_3 & x_4 \end{pmatrix}, \right.$$

$$P_4 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_1 & x_4 & x_3 \end{pmatrix}, \; P_5 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_3 & x_4 & x_1 & x_2 \end{pmatrix}, \; P_6 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_3 & x_4 & x_2 & x_1 \end{pmatrix},$$

$$\left. P_7 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_4 & x_3 & x_1 & x_2 \end{pmatrix}, \; P_8 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_4 & x_3 & x_2 & x_1 \end{pmatrix} \right\}.$$

The set $\{P_1,...P_8\}$ is the Galois group of the quartic in F. It is a subgroup of the full set of twentyfour permutations. We can form an infinite number of true relations among the roots in F. If we extend the field F by adjoining irrational expressions of the coefficients, other true relations among the roots can be formed in the extended field, F'. Consider, for example, the extended field formed by adjoining μ $(= (a_2^2 - 4a_4))$ to F so that the relation

$$x_1^2 - x_3^2 = \mu \text{ is in F'.}$$

We have met the relations

$$x_1 = -x_2 \text{ and } x_3 = -x_4$$

so that

$$x_1^2 = x_2^2 \text{ and } x_3^2 = x_4^2.$$

Another relation in F' is therefore

$$x_2^2 - x_4^2 = \mu.$$

The permutations that leave these relations true in F' are then

$\{P_1, P_2, P_3, P_4\}$.

This set is the Galois group of the quartic in F'. It is a subgroup of the set $\{P_1,...P_8\}$.

If we extend the field F' by adjoining the irrational expression $((-a_2 - \mu)/2)^{1/2}$ to form the field F" then the relation

$$x_3 - x_4 = 2((-a_2 - \mu)/2)^{1/2} \text{ is in F"}.$$

This relation is invariant under the two permutations

$\{P_1, P_3\}$.

This set is the Galois group of the quartic in F". It is a subgroup of the set $\{P_1, P_2, P_3, P_4\}$.

If, finally, we extend the field F" by adjoining the irrational $((-a_2 + \mu)/2)^{1/2}$ to form the field F"' then the relation

$$x_1 - x_2 = 2((-a_2 - \mu)/2)^{1/2} \text{ is in F"'}.$$

This relation is invariant under the identity transformation, P_1 , alone; it is the Galois group of the quartic in F".

The full group, and the subgroups, associated with the quartic equation are of order 24, 8, 4, 2, and 1. (The order of a group is the number of distinct elements that it contains). In **5.4**, we shall prove that the order of a subgroup is always an integral divisor of the order of the full group. The order of the full group divided by the order of a subgroup is called the index of the subgroup.

Galois introduced the idea of a normal or invariant subgroup: if H is a normal subgroup of G then

$$HG - GH = [H,G] = 0.$$

(H commutes with every element of G, see **5.5**).

Normal subgroups are also called either invariant or self-conjugate subgroups.

A normal subgroup H is *maximal* if no other subgroup of G contains H.

2.5. Solvability of polynomial equations

Galois defined the group of a given polynomial equation to be either the symmetric group, S_n, or a subgroup of S_n, (see **5.6**). The Galois method therefore involves the following steps:

1. The determination of the Galois group, Ga, of the equation.

2. The choice of a maximal subgroup of $H_{max(1)}$. In the above case, $\{P_1, ...P_8\}$ is a maximal subgroup of $Ga = S_4$.

3. The choice of a maximal subgroup of $H_{max(1)}$ from step 2.

In the above case, $\{P_1,..P_4\} = H_{max(2)}$ is a maximal subgroup of $H_{max(1)}$.

The process is continued until $H_{max} = \{P_1\} = \{I\}$.

The groups Ga, $H_{max(1)}$, ..,$H_{max(k)} = I$, form a *composition series*. The composition indices are given by the ratios of the successive orders of the groups:

$$g_n/h_{(1)},\ h_{(1)}/h_{(2)},\ ...h_{(k-1)}/1.$$

The composition indices of the symmetric groups S_n for n = 2 to 7 are found to be:

n	Composition Indices
2	2

3 2, 3

4 2, 3, 2, 2

5 2, 60

6 2, 360

7 2, 2520

We shall state, without proof, Galois' theorem:

A polynomial equation can be solved algebraically if and only if its group is solvable.

Galois defined a solvable group as one in which the composition indices are all prime numbers. Furthermore, he showed that if n > 4, the sequence of maximal normal subgroups is S_n, A_n, I, where A_n is the Alternating Group with (n!)/2 elements. The composition indices are then 2 and (n!)/2. For n > 4, however, (n!)/2 is not prime, therefore the groups S_n are not solvable for n > 4. Using Galois' Theorem, we see that it is therefore not possible to solve, algebraically, a general polynomial equation of degree n > 4.

3

SOME ALGEBRAIC INVARIANTS

Although algebraic invariants first appeared in the works of Lagrange and Gauss in connection with the Theory of Numbers, the study of algebraic invariants as an independent branch of Mathematics did not begin until the work of Boole in 1841. Before discussing this work, it will be convenient to introduce matrix versions of real bilinear forms, B, defined by

$$B = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j$$

where

$\mathbf{x} = [x_1, x_2, \ldots x_m]$, an m-vector,

$\mathbf{y} = [y_1, y_2, \ldots y_n]$, an n-vector,

and a_{ij} are real coefficients. The square brackets denote a *column* vector.

In matrix notation, the bilinear form is

$$B = \mathbf{x}^T \mathbf{A} \mathbf{y}$$

where

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

The scalar product of two n-vectors is seen to be a special case of a bilinear form in which $\mathbf{A} = \mathbf{I}$.

If $\mathbf{x} = \mathbf{y}$, the bilinear form becomes a quadratic form, Q:

$$Q = \mathbf{x}^T \mathbf{A} \mathbf{x}.$$

3.1. Invariants of binary quadratic forms

Boole began by considering the properties of the binary quadratic form

$$Q(x,y) = ax^2 + 2hxy + by^2$$

under a linear transformation of the coordinates

$$\mathbf{x}' = \mathbf{M}\mathbf{x}$$

where

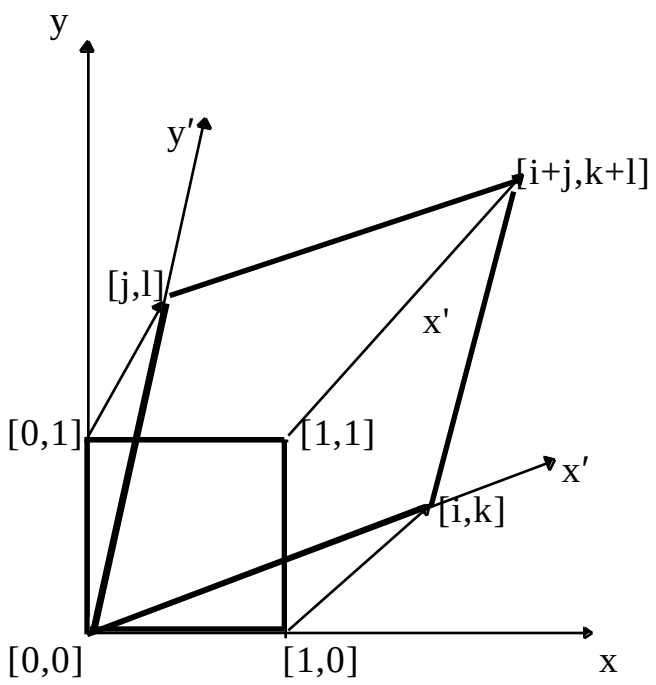
$$\mathbf{x} = [x,y],$$

$$\mathbf{x}' = [x',y'],$$

and

$$\mathbf{M} = \begin{pmatrix} i & j \\ k & l \end{pmatrix}.$$

The matrix **M** transforms an orthogonal coordinate system into an oblique coordinate system in which the new x'- axis has a slope (k/i), and the new y'- axis has a slope (l/j), as shown:



*The transformation of a unit square under **M**.*

The transformation is linear, therefore the new function Q'(x',y') is a binary quadratic:

$$Q'(x',y') = a'x'^2 + 2h'x'y' + b'y'^2.$$

The original function can be written

$$Q(x,y) = \mathbf{x}^T\mathbf{D}\mathbf{x}$$

where

$$\mathbf{D} = \begin{pmatrix} a & h \\ h & b \end{pmatrix},$$

and the determinant of **D** is

$$\det\mathbf{D} = ab - h^2, \text{ called the discriminant of Q.}$$

The transformed function can be written

$$Q'(x',y') = \mathbf{x}'^T\mathbf{D}'\mathbf{x}'$$

where

$$\mathbf{D}' = \begin{pmatrix} a' & h' \\ h' & b' \end{pmatrix},$$

and

$$\det\mathbf{D}' = a'b' - h'^2, \text{ the discriminant of Q'.}$$

Now,

$$\begin{aligned} Q'(x',y') &= (\mathbf{M}\mathbf{x})^T\mathbf{D}'\mathbf{M}\mathbf{x} \\ &= \mathbf{x}^T\mathbf{M}^T\mathbf{D}'\mathbf{M}\mathbf{x} \end{aligned}$$

and this is equal to Q(x,y) if

$$\mathbf{M}^T\mathbf{D}'\mathbf{M} = \mathbf{D}.$$

The invariance of the form Q(x,y) under the coordinate transformation **M** therefore leads to the relation

$$(\det\mathbf{M})^2 \det\mathbf{D}' = \det\mathbf{D}$$

because

$$\det\mathbf{M}^T = \det\mathbf{M}.$$

The explicit form of this equation involving determinants is

$$(il - jk)^2(a'b' - h'^2) = (ab - h^2).$$

The discriminant $(ab - h^2)$ of Q is said to be an *invariant* of the transformation because it is equal to the discriminant $(a'b' - h'^2)$ of Q', apart from a factor $(il - jk)^2$ that depends on the transformation itself, and not on the arguments a,b,h of the function Q.

3.2. General algebraic invariants

The study of general algebraic invariants is an important branch of Mathematics.

A binary form in two variables is

$$f(x_1,x_2) = a_o x_1^n + a_1 x_1^{n-1} x_2 + \ldots a_n x_2^n$$

$$= \sum a_i x_1^{n-i} x_2^i$$

If there are three or four variables, we speak of ternary forms or quaternary forms.

A binary form is transformed under the linear transformation **M** as follows

$$f(x_1,x_2) \Rightarrow f'(x_1',x_2') = a_o' x_1'^n + a_1' x_1'^{n-1} x_2' + \ldots$$

The coefficients

$$a_o, a_1, a_2,.. \neq a_o', a_1', a_2' ..$$

and the roots of the equation

$$f(x_1,x_2) = 0$$

differ from the roots of the equation

$$f'(x_1',x_2') = 0.$$

Any function $I(a_o,a_1,...a_n)$ of the coefficients of f that satisfies

$$r^w I(a_o',a_1',...a_n') = I(a_o,a_1,...a_n)$$

is said to be an *invariant* of f if the quantity r depends only on the transformation matrix **M**, and not on the coefficients a_i of the function being transformed. The degree of the invariant is the degree of the coefficients, and the exponent w is called the weight. In the example discussed above, the degree is two, and the weight is two.

Any function, C, of the coefficients *and* the variables of a form f that is invariant under the transformation **M**, except for a multiplicative factor that is a power of the discriminant of **M**, is said to be a covariant of f. For binary forms, C therefore satisfies

$$r^w C(a_o',a_1',...a_n'; x_1',x_2') = C(a_o,a_1,...a_n; x_1,x_2).$$

It is found that the Jacobian of two binary quadratic forms, $f(x_1,x_2)$ and $g(x_1,x_2)$, namely the determinant

$$\begin{vmatrix} \partial f/\partial x_1 & \partial f/\partial x_2 \\ \partial g/\partial x_1 & \partial g/\partial x_2 \end{vmatrix}$$

where $\partial f/\partial x_1$ is the partial derivative of f with respect to x_1 etc., is a simultaneous covariant of weight one of the two forms.

The determinant

$$\begin{vmatrix} \partial^2 f/\partial x_1^{\,2} & \partial^2 f/\partial x_1 \partial x_2 \\ \partial^2 g/\partial x_2 \partial x_1 & \partial^2 g/\partial x_2^{\,2} \end{vmatrix},$$

called the Hessian of the binary form f, is found to be a covariant of weight two. A full discussion of the general problem of algebraic invariants is outside the scope of this book. The following example will, however, illustrate the method of finding an invariant in a particular case.

Example:

To show that

$$(a_o a_2 - a_1^{\,2})(a_1 a_3 - a_2^{\,2}) - (a_o a_3 - a_1 a_2)^2/4$$

is an invariant of the binary cubic

$$f(x,y) = a_o x^3 + 3a_1 x^2 y + 3a_2 x y^2 + a_3 y^3$$

under a linear transformation of the coordinates.

The cubic may be written

$$f(x,y) = (a_o x^2 + 2a_1 xy + a_2 y^2)x + (a_1 x^2 + 2a_2 xy + a_3 y^2)y$$
$$= \mathbf{x}^T \mathbf{D} \mathbf{x}$$

where

$$\mathbf{x} = [x,y],$$

and

$$\mathbf{D} = \begin{pmatrix} a_o x + a_1 y & a_1 x + a_2 y \\ a_1 x + a_2 y & a_2 x + a_3 y \end{pmatrix}.$$

Let **x** be transformed to **x**': **x**' = **M**x, where

$$\mathbf{M} = \begin{pmatrix} i & j \\ k & l \end{pmatrix}$$

then

$$f(x,y) = f'(x',y')$$

if

$$\mathbf{D} = \mathbf{M}^T\mathbf{D}'\mathbf{M}.$$

Taking determinants, we obtain

$$\det\mathbf{D} = (\det\mathbf{M})^2\det\mathbf{D}',$$

an invariant of f(x,y) under the transformation **M**.

In this case, **D** is a function of x and y. To emphasize this point, put

$$\det\mathbf{D} = \varphi(x,y)$$

and

$$\det\mathbf{D}' = \varphi'(x',y')$$

so that

$$\varphi(x,y) = (\det\mathbf{M})^2\varphi'(x',y'$$

$$= (a_o x + a_1 y)(a_2 x + a_3 y) - (a_1 x + a_2 y)^2$$

$$= (a_o a_2 - a_1^2)x^2 + (a_o a_3 - a_1 a_2)xy + (a_1 a_3 - a_2^2)y^2$$

$$= \mathbf{x}^T\mathbf{E}\mathbf{x},$$

where

$$\mathbf{E} = \begin{pmatrix} (a_o a_2 - a_1^{\,2}) & (a_o a_3 - a_1 a_2)/2 \\ (a_o a_3 - a_1 a_2)/2 & (a_1 a_3 - a_2^{\,2}) \end{pmatrix}.$$

Also, we have

$$\varphi(x',y') = \mathbf{x}'^{T}\mathbf{E}'\mathbf{x}'$$

$$= \mathbf{x}^{T}\mathbf{M}^{T}\mathbf{E}'\mathbf{M}\mathbf{x}$$

therefore

$$\mathbf{x}^{T}\mathbf{E}\mathbf{x} = (\det\mathbf{M})^{2}\mathbf{x}^{T}\mathbf{M}^{T}\mathbf{E}'\mathbf{M}\mathbf{x}$$

so that

$$\mathbf{E} = (\det\mathbf{M})^{2}\mathbf{M}^{T}\mathbf{E}'\mathbf{M}.$$

Taking determinants, we obtain

$$\det\mathbf{E} = (\det\mathbf{M})^{4}\det\mathbf{E}'$$

$$= (a_o a_2 - a_1^{\,2})(a_1 a_3 - a_2^{\,2}) - (a_o a_3 - a_1 a_2)^2/4$$

= invariant of the binary cubic f(x,y) under the transformation $\mathbf{x}' = \mathbf{M}\mathbf{x}$.

4

SOME INVARIANTS OF PHYSICS

4.1. Galilean invariance.

Events of finite extension and duration are part of the physical world. It will be convenient to introduce the notion of *ideal events* that have neither extension nor duration. Ideal events may be represented as mathematical points in a space-time geometry. A particular event, **E**, is described by the four components [t,x,y,z] where t is the time of the event,

and x,y,z, are its three spatial coordinates. The time and space coordinates are referred to arbitrarily chosen origins. The spatial mesh need not be Cartesian.

Let an event **E**[t,x], recorded by an observer O at the origin of an x-axis, be recorded as the event **E**'[t',x'] by a second observer O', moving at constant speed V along the x-axis. We suppose that their clocks are synchronized at t = t' = 0 when they coincide at a common origin, x = x' = 0.

At time t, we write the plausible equations

$$t' = t$$

and

$$x' = x - Vt,$$

where Vt is the distance travelled by O' in a time t. These equations can be written

$$\mathbf{E}' = \mathbf{G}\mathbf{E}$$

where

$$\mathbf{G} = \begin{pmatrix} 1 & 0 \\ -V & 1 \end{pmatrix}$$

G is the operator of the Galilean transformation.

The inverse equations are

$$t = t'$$

and

$$x = x' + Vt'$$

or

$$\mathbf{E} = \mathbf{G}^{-1}\mathbf{E}'$$

where $\mathbf{G}^{-1}$ is the inverse Galilean operator. (It undoes the effect of **G**).

If we multiply t and t' by the constants k and k', respectively, where k and k' have dimensions of velocity then all terms have dimensions of length.

In space-space, we have the Pythagorean form $x^2 + y^2 = r^2$, an invariant under rotations. We are therefore led to ask the question: is $(kt)^2 + x^2$ invariant under the operator **G** in space-time? Calculation gives

$$(kt)^2 + x^2 = (k't')^2 + x'^2 + 2Vx't' + V^2t'^2$$

$$= (k't')^2 + x'^2 \text{ only if } V = 0.$$

We see, therefore, that Galilean space-time is not Pythagorean in its algebraic form. We note, however, the key role played by *acceleration* in Galilean-Newtonian physics:

The velocities of the events according to O and O' are obtained by differentiating the equation $x' = -Vt + x$ with respect to time, giving

$$v' = -V + v,$$

a plausible result, based upon our experience.

Differentiating v' with respect to time gives

$$dv'/dt' = a' = dv/dt = a$$

where a and a' are the accelerations in the two frames of reference. The classical acceleration is *invariant* under the Galilean transformation. If the relationship $v' = v - V$ is used to describe the motion of a pulse of light,

moving in empty space at v = c 3 x 10^8 m/s, it does not fit the facts. All studies of very high speed particles that emit electromagnetic radiation show that v' = c for *all* values of the relative speed, V.

4.2. Lorentz invariance and Einstein's space-time symmetry.

It was Einstein, above all others, who advanced our understanding of the true nature of space-time and relative motion. We shall see that he made use of a symmetry argument to find the changes that must be made to the Galilean transformation if it is to account for the relative motion of rapidly moving objects and of beams of light. He recognized an inconsistency in the Galilean-Newtonian equations, based as they are, on everyday experience. Here, we shall restrict the discussion to non-accelerating, or so called inertial, frames

We have seen that the classical equations relating the events **E** and **E'** are **E'** = **GE**, and the inverse **E** = $\mathbf{G}^{-1}\mathbf{E}'$

where

$$\mathbf{G} = \begin{pmatrix} 1 & 0 \\ -V & 1 \end{pmatrix} \text{ and } \mathbf{G}^{-1} = \begin{pmatrix} 1 & 0 \\ V & 1 \end{pmatrix}$$

These equations are connected by the substitution $V \leftrightarrow -V$; this is an algebraic statement of the Newtonian principle of relativity. Einstein incorporated this principle in his theory. He also retained the *linearity* of the classical equations in the absence of any evidence to the contrary.

(Equispaced intervals of time and distance in one inertial frame remain equispaced in any other inertial frame). He therefore *symmetrized* the space-time equations as follows:

$$\begin{pmatrix} t' \\ x' \end{pmatrix} = \begin{pmatrix} 1 & -V \\ -V & 1 \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} \quad .$$

Note, however, the inconsistency in the dimensions of the time-equation that has now been introduced:

$$t' = t - Vx.$$

The term Vx has dimensions of $[L]^2/[T]$, and not [T]. This can be corrected by introducing the invariant speed of light, c ¾ a postulate in Einstein's theory that is consistent with experiment:

$$ct' = ct - Vx/c$$

so that all terms now have dimensions of length.

Einstein went further, and introduced a dimensionless quantity γ instead of the scaling factor of unity that appears in the Galilean equations of space-time. This factor must be consistent with all observations. The equations then become

$$ct' = \gamma ct - \beta\gamma x$$

$$x' = -\beta\gamma ct + \gamma x, \text{ where } \beta = V/c.$$

These can be written

$$\mathbf{E}' = \mathbf{L}\mathbf{E},$$

where

$$\mathbf{L} = \begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix}, \text{ and } \mathbf{E} = [ct, x]$$

L is the operator of the Lorentz transformation.

The inverse equation is

$$\mathbf{E} = \mathbf{L}^{-1}\mathbf{E}'$$

where

$$\mathbf{L}^{-1} = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} .$$

This is the inverse Lorentz transformation, obtained from **L** by changing $\beta \rightarrow -\beta$ (or ,$V \rightarrow -V$); it has the effect of undoing the transformation **L**. We can therefore write

$$\mathbf{L}\mathbf{L}^{-1} = \mathbf{I}$$

or

$$\begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Equating elements gives

$$\gamma^2 - \beta^2\gamma^2 = 1$$

therefore,

$$\gamma = 1/\sqrt{(1 - \beta^2)} \text{ (taking the positive root).}$$

4.3. The invariant interval.

Previously, it was shown that the space-time of Galileo and Newton is not Pythagorean in form. We now ask the question: is Einsteinian space-time Pythagorean in form? Direct calculation leads to

$$(ct)^2 + (x)^2 = \gamma^2(1 + \beta^2)(ct')^2 + 4\beta\gamma^2 x'ct'$$
$$+\gamma^2(1 + \beta^2)x'^2$$
$$\neq (ct')^2 + (x')^2 \text{ if } \beta > 0.$$

Note, however, that the *difference of squares* is an invariant under **L**:

$$(ct)^2 - (x)^2 = (ct')^2 - (x')^2$$

because

$$\gamma^2(1 - \beta^2) = 1.$$

Space-time is said to be pseudo-Euclidean.

The negative sign that characterizes Lorentz invariance can be included in the theory in a general way as follows.

We introduce two kinds of 4-vectors

$$x^\mu = [x^0, x^1, x^2, x^3], \text{ a } \textit{contravariant} \text{ vector,}$$

and

$$x_\mu = [x_0, x_1, x_2, x_3], \text{ a } \textit{covariant} \text{ vector, where}$$
$$x_\mu = [x^0, -x^1, -x^2, -x^3].$$

The scalar product of the vectors is defined as

$$x^{\mu T}x_\mu = (x^0, x^1, x^2, x^3)[x^0, -x^1, -x^2, -x^3]$$

$$= (x^0)^2 - ((x^1)^2 + (x^2)^2 + (x^3)^2)$$

The event 4-vector is

$E^\mu = [ct, x, y, z]$ and the covariant form is

$E_\mu = [ct, -x, -y, -z]$

so that the Lorentz invariant scalar product is

$$E^{\mu T}E_\mu = (ct)^2 - (x^2 + y^2 + z^2).$$

The 4-vector x^μ transforms as follows:

$$x'^\mu = \mathbf{L}x^\mu$$

where

$$\mathbf{L} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

This is the operator of the Lorentz transformation if the motion of O' is along the x-axis of O's frame of reference.

Important consequences of the Lorentz transformation are that intervals of time measured in two different inertial frames are not the same but are related by the equation

$$\Delta t' = \gamma \Delta t$$

where Δt is an interval measured on a clock at rest in O's frame, and distances are given by

$$\Delta l' = \Delta l / \gamma$$

where Δl is a length measured on a ruler at rest in O's frame.

4.4. The energy-momentum invariant.

A differential time interval, dt, cannot be used in a Lorentz-invariant way in kinematics. We must use the proper time differential interval, dτ, defined by

$$(cdt)^2 - dx^2 = (cdt')^2 - dx'^2 \equiv (cd\tau)^2.$$

The Newtonian 3-velocity is

$$\mathbf{v}_N = [dx/dt,\ dy/dt,\ dz/dt],$$

and this must be replaced by the 4-velocity

$$\begin{aligned} V^\mu &= [d(ct)/d\tau,\ dx/d\tau,\ dy/d\tau,\ dz/d\tau] \\ &= [d(ct)/dt,\ dx/dt,\ dy/dt,\ dz/dt]dt/d\tau \\ &= [\gamma c, \gamma \mathbf{v}_N]\ . \end{aligned}$$

The scalar product is then

$$\begin{aligned} V^\mu V_\mu &= (\gamma c)^2 - (\gamma v_N)^2 \\ &= (\gamma c)^2(1 - (v_N/c)^2) \\ &= c^2. \end{aligned}$$

(In forming the scalar product, the transpose is understood).

The magnitude of the 4-velocity is $\|V^\mu\| = c$, the invariant speed of light.

In Classical Mechanics, the concept of momentum is important because of its role as an invariant in an isolated system. We therefore introduce the concept of 4-momentum in Relativistic Mechanics in order to find possible Lorentz invariants involving this new quantity. The contravariant 4-momentum is defined as:

$$P^{\mu} = mV^{\mu}$$

where m is the mass of the particle. (It is a Lorentz scalar, the mass measured in the frame in which the particle is at rest).

The scalar product is

$$P^{\mu}P_{\mu} = (mc)^2.$$

Now,

$$P^{\mu} = [m\gamma c,\ m\gamma \mathbf{v}_N]$$

therefore,

$$P^{\mu}P_{\mu} = (m\gamma c)^2 - (m\gamma v_N)^2.$$

Writing

$M = \gamma m$, the relativistic mass, we obtain

$$P^{\mu}P_{\mu} = (Mc)^2 - (Mv_N)^2 = (mc)^2.$$

Multiplying throughout by c^2 gives

$$M^2c^4 - M^2v_N{}^2c^2 = m^2c^4.$$

The quantity Mc^2 has dimensions of energy; we therefore write

$$E = Mc^2$$

the total energy of a freely moving particle.

This leads to the *fundamental invariant of dynamics*

$$c^2P^{\mu}P_{\mu} = E^2 - (pc)^2 = E^{o2}$$

where

$E^o = mc^2$ is the rest energy of the particle, and **p** is its *relativistic* 3-momentum.

The total energy can be written:

$$E = \gamma E^o = E^o + T,$$

where

$$T = E^o(\gamma - 1),$$

the relativistic kinetic energy.

The magnitude of the 4-momentum is a Lorentz invariant

$$\|P^\mu\| = mc.$$

The 4- momentum transforms as follows:

$$P'^\mu = \mathbf{L}P^\mu.$$

For relative motion along the x-axis, this equation is equivalent to the equations

$$E' = \gamma E - \beta\gamma cp^x$$

and

$$cp^x = -\beta\gamma E + \gamma cp^x .$$

Using the Planck-Einstein equations $E = h\nu$ and

$E = p^x c$ for photons, the energy equation becomes

$$\begin{aligned} \nu' &= \gamma\nu - \beta\gamma\nu \\ &= \gamma\nu(1 - \beta) \\ &= \nu(1 - \beta)/(1 - \beta^2)^{1/2} \\ &= \nu[(1 - \beta)/(1 + \beta)]^{1/2} . \end{aligned}$$

This is the relativistic Doppler shift for the frequency ν', measured in an inertial frame (primed) in terms of the frequency ν measured in another inertial frame (unprimed).

4.5. The frequency-wavenumber invariant

Particle-Wave duality, one of the most profound discoveries in Physics, has its origins in Lorentz invariance. It was proposed by deBroglie in the early 1920's. He used the following argument.

The displacement of a wave can be written

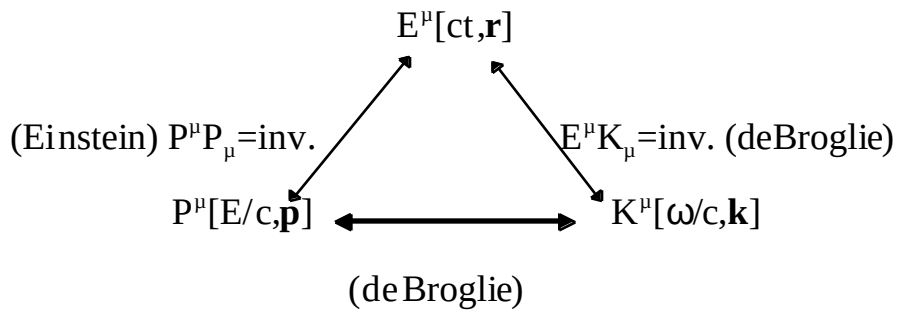
$$y(t,\mathbf{r}) = A\cos(\omega t - \mathbf{k}\cdot\mathbf{r})$$

where $\omega = 2\pi\nu$ (the angular frequency), $\|\mathbf{k}\| = 2\pi/\lambda$ (the wavenumber), and $\mathbf{r} = [x, y, z]$ (the position vector). The phase $(\omega t - \mathbf{k}\cdot\mathbf{r})$ can be written $((\omega/c)ct - \mathbf{k}\cdot\mathbf{r})$, and this has the form of a Lorentz invariant obtained from the 4-vectors

$$E^{\mu}[ct, \mathbf{r}], \text{ and } K^{\mu}[\omega/c, \mathbf{k}]$$

where E^{μ} is the event 4-vector, and K^{μ} is the "frequency-wavenumber" 4-vector.

deBroglie noted that the 4-momentum P^{μ} is connected to the event 4-vector E^{μ} through the 4-velocity V^{μ}, and the frequency-wavenumber 4-vector K^{μ} is connected to the event 4-vector E^{μ} through the Lorentz invariant phase of a wave $((\omega/c)ct - \mathbf{k}\bullet\mathbf{r})$. He therefore proposed that a direct connection must exist between P^{μ} and K^{μ}; it is illustrated in the following diagram:



The coupling between P^μ and K^μ via E^μ.

deBroglie proposed that the connection is the simplest possible, namely, P^μ and K^μ are proportional to each other. He realized that there could be only one value for the constant of proportionality if the Planck-Einstein result for photons $E = h\omega/2\pi$ is but a special case of a general result, it must be $h/2\pi$, where h is Planck's constant. Therefore, deBroglie proposed that

$$P^\mu \quad K^\mu$$

or

$$P^\mu = (h/2\pi)K^\mu.$$

Equating the elements of the 4-vectors gives

$$E = (h/2\pi)\omega$$

and

$$\mathbf{p} = (h/2\pi)\mathbf{k}\ .$$

In these remarkable equations, our notions of particles and waves are forever merged. The smallness of the value of Planck's constant prevents

us from observing the duality directly; however, it is clearly observed at the molecular, atomic, nuclear, and particle level.

4.6. deBroglie's invariant.

The invariant formed from the frequency-wavenumber 4-vector is

$$K^{\mu}K_{\mu} = (\omega/c, \mathbf{k})[\omega/c, -\mathbf{k}]$$

$$= (\omega/c)^2 - k^2 = (\omega^o/c)^2,$$ where ω^o is the proper angular frequency.

This invariant is the wave version of Einstein's energy-momentum invariant; it gives the dispersion relation

$$\omega^{o2} = \omega^2 - (kc)^2.$$

The ratio ω/k is the phase velocity of the wave, v_{φ}.

For a wave-packet, the group velocity v_G is $d\omega/dk$; it can be obtained by differentiating the dispersion equation as follows:

$$\omega d\omega - kc^2 dk = 0$$

therefore,

$$v_G = d\omega/dk = kc^2/\omega.$$

The deBroglie invariant involving the product of the phase and group velocity is therefore

$$v_{\varphi}v_G = (\omega/k).(kc^2/\omega) = c^2.$$

This is the wave-equivalent of Einstein's famous

$$E = Mc^2.$$

We see that

$$v_\varphi v_G = c^2 = E/M$$

or,

$v_G = E/Mv_\varphi = Ek/M\omega = p/M = v_N$, the particle velocity.

This result played an important part in the development of Wave Mechanics.

We shall find in later chapters, that Lorentz transformations form a group, and that invariance principles are related directly to symmetry transformations and their associated groups.

5

GROUPS — CONCRETE AND ABSTRACT

5.1 Some concrete examples

The elements of the set $\{\pm 1, \pm i\}$, where $i = \sqrt{-1}$, are the roots of the equation $x^4 = 1$, the "fourth roots of unity". They have the following special properties:

1. The product of any two elements of the set (including the same two elements) is always an element of the set. (The elements obey *closure*).

2. The order of combining pairs in the triple product of any elements of the set does not matter. (The elements obey *associativity*).

3. A unique element of the set exists such that the product of any element of the set and the unique element (called the *identity*) is equal to the element itself. (An *identity* element exists).

4. For each element of the set, a corresponding element exists such that the product of the element and its corresponding element (called the inverse) is equal to the identity. (An *inverse* element exists).

The set of elements {±1, ±i} with these four properties is said to form a *GROUP*.

In this example, the *law of composition* of the group is multiplication; this need not be the case. For example, the set of integers Z = {.., −2, −1, 0, 1, 2, ...} forms a group if the law of composition is *addition*. In this group, the identity element is zero, and the inverse of each integer is the integer with the same magnitude but with opposite sign.

In a different vein, we consider the set of 4×4 matrices:

$$\{M\} = \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix}, \begin{pmatrix} 0&0&0&1 \\ 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \end{pmatrix}, \begin{pmatrix} 0&0&1&0 \\ 0&0&0&1 \\ 1&0&0&0 \\ 0&1&0&0 \end{pmatrix}, \begin{pmatrix} 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \\ 1&0&0&0 \end{pmatrix}.$$

If the law of composition is matrix multiplication , then {M} is found to obey:

1 --closure

and

2 --associativity,

and to contain:

3 --an identity, diag(1, 1, 1, 1),

and

4 --inverses.

The set {M} forms a group under matrix multilication.

5.2. Abstract groups

The examples given above illustrate the generality of the group concept. In the first example, the group elements are real and imaginary numbers, in the second, they are positive and negative integers, and in the third, they are matrices that represent linear operators (see later discussion). Cayley, in the mid-19th century, first emphasized this generality, and he introduced the concept of an *abstract group*, G_n which is a collection of n distinct elements $(...g_i...)$ for which a law of composition is given. If n is finite, the group is said to be a *group of order n*. The collection of elements must obey the four rules:

1. If g_i, g_j G then $g_n = g_j \cdot g_i$ G g_i, g_j G (closure)

2. $g_k(g_jg_i) = (g_kg_j)g_i$ [leaving out the composition symbol $\cdot$] (associativity)

3. e G such that $g_ie = eg_i = g_i$ g_i G (an identity exists)

4. If g_i G then g_i^{-1} G such that $g_i^{-1}g_i = g_ig_i^{-1} = e$ (an inverse exists).

For finite groups, the group structure is given by listing all compositions of pairs of elements in a *group table*, as follows:

	e . g_i g_j . ← (1st symbol, or operation, in pair)
e	. . . .
.	. . . .
g_i	. g_ig_i g_ig_j .
g_j	. g_jg_i g_jgj .
g_k	. g_kg_i g_kg_j .
.	
.	

If $g_j g_i = g_i g_j \quad g_i, g_j \quad G$, then G is said to be a *commutative* or *abelian* group. The group table of an abelian group is symmetric under reflection in the diagonal.

A group of elements that has the same structure as an abstract group is a *realization* of the group.

5.3 The dihedral group, D_3

The set of operations that leaves an equilateral triangle invariant under rotations in the plane about its center, and under reflections in the three planes through the vertices, perpendicular to the opposite sides, forms a group of six elements. A study of the structure of this group (called the dihedral group, D_3) illustrates the typical group-theoretical approach.

The geometric operations that leave the triangle invariant are:

Rotations about the z-axis (anticlockwise rotations are positive)

$R_z(0) \quad (123) \rightarrow (123) = e$, the identity

$R_z(2\pi/3)(123) \rightarrow (312) = a$

$R_z(4\pi/3)(123) \rightarrow (231) = a^2$

and reflections in the planes I, II, and III:

$R_I\ (123) \rightarrow (123) = b$

$R_{II}\ (123) \rightarrow (321) = c$

$R_{III}\ (123) \rightarrow (213) = d$

This set of operators is $D_3 = \{e, a, a^2, b, c, d\}$.

Positive rotations are in an anticlockwise sense and the inverse rotations are in a clockwise sense., so that the inverse of e, a, a^2 are

$e^{-1} = e$, $a^{-1} = a^2$, and $(a^2)^{-1} = a$.

The inverses of the reflection operators are the operators themselves:

$b^{-1} = b$, $c^{-1} = c$, and $d^{-1} = d$.

We therefore see that the set D_3 forms a group. The group multiplication table is:

	e	a	a^2	b	c	d
e	e	a	a^2	b	c	d
a	a	a^2	e	d	b	c
a^2	a^2	e	a	c	d	b
b	b	c	d	e	a	a^2
c	c	d	b	a^2	e	a
d	d	b	c	a	a^2	e

In reading the table, we follow the rule that the first operation is written on the right: for example, $ca^2 = b$. A feature of the group D_3 is that it can be subdivided into sets of either rotations involving $\{e, a, a^2\}$ or reflections involving $\{b, c, d\}$. The set $\{e, a, a^2\}$ forms a group called the *cyclic group* of order three, C_3. A group is cyclic if all the elements of the group are powers of a single element. The cyclic group of order n, C_n, is

$$C_n = \{e, a, a^2, a^3, ...,a^{n-1}\},$$

where n is the smallest integer such that $a^n = e$, the identity. Since

$$a^k a^{n-k} = a^n = e,$$

an inverse a^{n-k} exists. All cyclic groups are abelian.

The group D_3 can be broken down into a part that is a group C_3, and a part that is the product of one of the remaining elements and the elements of C_3. For example, we can write

$$D_3 = C_3 + bC_3 \text{ , } b \quad C_3$$

$$= \{e, a, a^2\} + \{b, ba, ba^2\}$$

$$= \{e, a, a^2\} + \{b, c, d\}$$

$$= cC_3 = dC_3.$$

This decomposition is a special case of an important theorem known as Lagrange's theorem. (Lagrange had considered permutations of roots of equations before Cauchy and Galois).

5.4 Lagrange's theorem

The order m of a subgroup H_m of a finite group G_n of order n is a factor (an integral divisor) of n.

Let

$G_n = \{g_1=e, g_2, g_3, ...g_n\}$ be a group of order n, and let

$H_m = \{h_1=e, h_2, h_3, ...h_m\}$ be a subgroup of G_n of order m.

If we take any element g_k of G_n which is not in H_m, we can form the set of elements

$$\{g_kh_1, g_kh_2, g_kh_3, ...g_kh_m\} \equiv g_kH_m.$$

This is called the *left coset* of H_m with respect to g_k. We note the important facts that all the elements of g_kh_j, j=1 to m are distinct, and that none of the elements g_kh_j belongs to H_m.

Every element g_k that belongs to G_n but does not belong to H_m belongs to some coset g_kH_m so that G_n forms the union of H_m and a number

of distinct (non-overlapping) cosets. (There are (n − m) such distinct cosets). Each coset has m different elements and therefore the order n of G_n is divisible by m, hence n = Km, where the integer K is called the *index* of the subgroup H_m under the group G_n. We therefore write

$$G_n = g_1H_m + g_{j2}H_m + g_{k3}H_m + ...g_{\nu K}H_m$$

where

g_{j2} G_n H_m,

g_{k3} G_n H_m, $g_{j2}H_m$

.

g_{nK} G_n H_m, $g_{j2}H_m$, $g_{k3}H_m$, ...$g_{n-1,\,K-1}H_m$.

The subscripts 2, 3, 4, ..K are the *indices* of the group.

As an example, consider the permutations of three objects 1, 2, 3 (the group S_3) and let $H_m = C_3 = \{123, 312, 231\}$, the cyclic group of order three. The elements of S_3 that are not in H_3 are {132, 213, 321}. Choosing $g_k = 132$, we obtain

$$g_kH_3 = \{132, 321, 213\},$$

and therefore

$$S_3 = C_3 + g_{k2}C_3 \text{ ,K = 2.}$$

This is the result obtained in the decomposition of the group D_3 , if we make the substitutions e = 123, a = 312, a^2 = 231, b = 132, c = 321, and d = 213.

The groups D_3 and S_3 are said to be *isomorphic.* Isomorphic groups have the same group multiplication table. Isomorphism is a special case of homomorphism that involves a many-to-one correspondence.

5.5 Conjugate classes and invariant subgroups

If there exists an element v G_n such that two elements a, b G_n are related by $vav^{-1} = b$, then b is said to be *conjugate* to a. A finite group can be separated into sets that are conjugate to each other.

The *class* of G_n is defined as the set of conjugates of an element a G_n. The element itself belongs to this set. If a is conjugate to b, the class conjugate to a and the class conjugate to b are the same. If a is not conjugate to b, these classes have no common elements. G_n can be decomposed into classes because each element of G_n belongs to a class.

An element of G_n that commutes with all elements of G_n forms a class by itself.

The elements of an abelian group are such that

$$bab^{-1} = a \text{ for all a, b } G_n,$$

so that

$$ba = ab.$$

If H_m is a subgroup of G_n, we can form the set

$$\{aea^{-1}, ah_2a^{-1}, ...ah_ma^{-1}\} = aH_ma^{-1} \text{ where a } G_n .$$

Now, aH_ma^{-1} is another subgroup of H_m in G_n. Different subgroups may be found by choosing different elements a of G_n. If, for all values of a G_n

$$aH_ma^{-1} = H_m$$

(all conjugate subgroups of H_m in G_n are identical to H_m),

then H_m is said to be an *invariant subgroup* in G_n.

Alternatively, H_m is an invariant in G_n if the left- and right-cosets formed with any a G_n are equal, i. e. $ah_i = h_ka$.

An invariant subgroup H_m of G_n commutes with all elements of G_n. Furthermore, if h_i H_m then all elements ah_ia^{-1} H_m so that H_m is an invariant subgroup of G_n if it contains elements of G_n in complete classes.

Every group G_n contains two trivial invariant subgroups, $H_m = G_n$ and $H_m = e$. A group with no proper (non-trivail) invariant subgroups is said to be *simple* (atomic). If none of the proper invariant subgroups of a group is abelian, the group is said to be *semisimple*.

An invariant subgroup H_m and its cosets form a group under multiplication called the *factor group* (written G_n/H_m) of H_m in G_n.

These formal aspects of Group Theory can be illustrated by considering the following example:

The group $D_3 = \{e, a, a^2, b, c, d\} \sim S_3 = \{123, 312, 231, 132, 321, 213\}$.

C_3 is a subgroup of S_3: $C_3 = H_3 = \{e, a, a^2\} = \{123, 312, 231\}$.

Now,

$$bH_3 = \{132, 321, 213\} = H_3b$$

$$cH_3 = \{321, 213, 132\} = H_3c$$

and

$$dH_3 = \{213,132, 321\} = H_3d.$$

Since H_3 is a proper invariant subgroup of S_3, we see that S_3 is not simple. H_3 is abelian therefore S_3 is not semisimple.

The decomposition of S_3 is

$$S_3 = H_3 + bH_3 = H_3 + H_3b.$$

and, in this case we have

$$H_3b = H_3c = H_3d.$$

(Since the index of H_3 is 2, H_3 must be invariant).

The conjugate classes are

$$e = e$$

$$eae^{-1} = ea = a$$

$$aaa^{-1} = ae = a$$

$$a^2a(a^2)^{-1} = a^2a^2 = a$$

$$bab^{-1} = bab = a^2$$

$$cac^{-1} = cac = a^2$$

$$dad^{-1} = dad = a^2$$

The class conjugate to a is therefore $\{a, a^2\}$.

The class conjugate to b is found to be $\{b, c, d\}$. The group S_3 can be decomposed by classes:

$$S_3 = \{e\} + \{a, a^2\} + \{b, c, d\}.$$

S_3 contains three conjugate classes.

If we now consider $H_m = \{e, b\}$ an abelian subgroup, we find

$$aH_m = \{a,d\},\ H_m a = \{a.c\},$$

$$a^2H_m = \{a^2,c\},\ H_m a^2 = \{a^2, d\},\ \text{etc.}$$

All left and right cosets are not equal: $H_m = \{e, b\}$ is therefore not an invariant subgroup of S_3. We can therefore write

$$S_3 = \{e, b\} + \{a, d\} + \{a^2, c\}$$

$$= H_m + aH_m + a^2H_m.$$

Applying Lagrange's theorem to S_3 gives the orders of the possible subgroups: they are

order 1: $\{e\}$

order 2: $\{e, d\}$; $\{e, c\}$: $\{e, d\}$

order 3: $\{e, a, a^2\}$ (abelian and invariant)

order 6: S_3.

5.6 Permutations

A permutation of the set $\{1, 2, 3, ...,n\}$ of n distinct elements is an ordered arrangement of the n elements. If the order is changed then the

permutation is changed. The number of permutations of n distinct elements is n!

We begin with a familiar example: the permutations of three distinct objects labelled 1, 2, 3. There are six possible arrangements; they are

123, 312, 231, 132, 321, 213.

These arrangements can be written conveniently in matrix form:

$$\pi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \pi_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \pi_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

$$\pi_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \pi_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \pi_6 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

The product of two permutations is the result of performing one arrangement after another. We then find

$$\pi_2\pi_3 = \pi_1$$

and

$$\pi_3\pi_2 = \pi_1$$

whereas

$$\pi_4\pi_5 = \pi_3$$

and

$$\pi_5\pi_4 = \pi_2.$$

The permutations π_1, π_2, π_3 commute in pairs (they correspond to the rotations of the dihedral group) whereas the permutations do not commute (they correspond to the reflections).

A general product of permutations can be written

$$\begin{pmatrix} s_1 & s_2 & . & . & .s_n \\ t_1 & t_2 & . & . & .t_n \end{pmatrix} \begin{pmatrix} 1 & 2 & . & . & n \\ s_1 & s_2 & . & . & s_n \end{pmatrix} = \begin{pmatrix} 1 & 2 & . & . & n \\ t_1 & t_2 & . & . & t_n \end{pmatrix}.$$

The permutations are found to have the following properties:

1. The product of two permutations of the set {1, 2, 3, ...} is itself a permutation of the set. (Closure)

2. The product obeys associativity:

$(\pi_k\pi_j)\pi_i = \pi_k(\pi_j\pi_i)$, (not generally commutative).

3. An identity permutation exists.

4. An inverse permutation exists:

$$\pi^{-1} = \begin{pmatrix} s_1 & s_2 & . & . & . & s_n \\ 1 & 2 & . & . & . & n \end{pmatrix}$$

such that $\pi\pi^{-1} = \pi^{-1}\pi$ = identity permutation.

The set of permutations therefore forms a group

5.7 Cayley's theorem:

Every finite group is isomorphic to a certain permutation group.

Let $G_n = \{g_1, g_2, g_3, \ . \ . \ . g_n\}$ be a finite group of order n. We choose any element g_i in G_n, and we form the products that belong to G_n:

$g_ig_1, g_ig_2, g_ig_3, \ . \ . \ . \ g_ig_n$.

These products are the n-elements of G_n rearranged. The permutation π_i, associated with g_i is therefore

$$\pi_i = \begin{pmatrix} g_1 & g_2 & . & . & g_n \\ g_ig_1 & g_ig_2 & . & . & g_ig_n \end{pmatrix}.$$

If the permutation π_j associated with g_j is

$$\pi_j = \begin{pmatrix} g_1 & g_2 & . & . & g_n \\ g_j g_1 & g_j g_2 & . & . & g_j g_n \end{pmatrix},$$

where $g_i \neq g_j$, then

$$\pi_j \pi_i = \begin{pmatrix} g_1 & g_2 & . & . & g_n \\ (g_j g_i) g_i & (g_j g_i) g_2 & . & . & (g_j g_i) g_n \end{pmatrix} .$$

This is the permutation that corresponds to the element $g_j g_i$ of G_n.

There is a direct correspondence between the elements of G_n and the n-permutations $\{\pi_1, \pi_2, . \ . \ . \pi_n\}$. The group of permutations is a subgroup of the full symmetric group of order n! that contains all the permutations of the elements $g_1, g_2, . \ . \ g_n$.

Cayley's theorem is important not only in the theory of finite groups but also in those quantum systems in which the *indistinguishability* of the fundamental particles means that certain quantities must be invariant under the exchange or permutation of the particles.

6

LIE'S DIFFERENTIAL EQUATION, INFINITESIMAL ROTATIONS AND ANGULAR MOMENTUM OPERATORS

Although the field of continuous transformation groups (Lie groups) has its origin in the theory of differential equations, we shall introduce the subject using geometrical ideas.

6.1 Coordinate and vector rotations

A 3-vector $\mathbf{v} = [v_x, v_y, v_z]$ transforms into $\mathbf{v}' = [v_x', v_y', v_z']$ under a general coordinate rotation R about the origin of an orthogonal coordinate system as follows:

$$\mathbf{v}' = R\ \mathbf{v},$$

where

$$R = \begin{pmatrix} \mathbf{i}.\mathbf{i}' & \mathbf{j}.\mathbf{i}' & \mathbf{k}.\mathbf{i}' \\ \mathbf{i}.\mathbf{j}' & \mathbf{j}.\mathbf{j}' & \mathbf{k}.\mathbf{j}' \\ \mathbf{i}.\mathbf{k}' & \mathbf{j}.\mathbf{k}' & \mathbf{k}.\mathbf{k}' \end{pmatrix}$$

$$= \begin{pmatrix} \cos\theta_{ii'} & . & . \\ \cos\theta_{ij'} & . & . \\ \cos\theta_{ik'} & . & \cos\theta_{kk'} \end{pmatrix}$$

where $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$ are orthogonal unit vectors, along the axes, before and after the transformation, and the $\cos\theta_{ii'}$'s are direction cosines.

The simplest case involves rotations in the x-y plane:

$$\begin{pmatrix} v_{x'} \\ v_{y'} \end{pmatrix} = \begin{pmatrix} \cos\theta_{ii'} & \cos\theta_{ji} \\ \cos\theta_{ij'} & \cos\theta_{jj'} \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

$$= \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix} = R_c(\varphi)\mathbf{v}$$

where $R_c(\varphi)$ is the coordinate rotation operator. If the vector is rotated in a fixed coordinate system, we have $\varphi \rightarrow -\varphi$ so that

$$\mathbf{v}' = R_v(\varphi)\mathbf{v},$$

where

$$R_v(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

6.2 Lie's differential equation

The main features of Lie's *Theory of Continuous Transformation Groups* can best be introduced by discussing the properties of the rotation operator $R_v(\varphi)$ when the angle of rotation is an infinitesimal. In general, $R_v(\varphi)$ transforms a point **P**[x, y] in the plane into a "new" point **P**′[x′, y′]: $\mathbf{P}' = R_v(\varphi)\mathbf{P}$. Let the angle of rotation be sufficiently small for us to put $\cos(\varphi) \approx 1$ and $\sin(\varphi) \approx \delta\varphi$, in which case, we have

$$R_v(\delta\varphi) = \begin{pmatrix} 1 & -\delta\varphi \\ \delta\varphi & 1 \end{pmatrix}$$

and

$$x' = x.1 - y\delta\varphi = x - y\delta\varphi$$

$$y' = x\delta\varphi + y.1 = x\delta\varphi + y$$

Let the corresponding changes $x \rightarrow x'$ and $y \rightarrow y'$ be written

$$x' = x + \delta x \text{ and } y' = y + \delta y$$

so that

$$\delta x = -y\delta\varphi \text{ and } \delta y = x\delta\varphi$$

We note that

$$R_v(\delta\varphi) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\delta\varphi$$

$$= \mathbf{I} + \mathbf{i}\delta\varphi$$

where

$$\mathbf{i} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = R_v(\pi/2).$$

Lie introduced another important way to interpret the operator $\mathbf{i} = R_v(\pi/2)$, that involves the *derivative* of $R_v(\varphi)$ evaluated at the identity value of the parameter, $\varphi = 0$:

$$dR_v(\varphi)/d\varphi\Big|_{\varphi=0} = \begin{pmatrix} -\sin\varphi & -\cos\varphi \\ \cos\varphi & -\sin\varphi \end{pmatrix}\Bigg|_{\varphi=0} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \mathbf{i}$$

so that

$$R_v(\delta\varphi) = \mathbf{I} + dR_v(\varphi)/d\varphi\Big|_{\varphi=0}\delta\varphi,$$

a quantity that differs from the identity **I** by a term that involves the infinitesimal, $\delta\varphi$: *this is an infinitesimal transformation*.

Lie was concerned with Differential Equations and not Geometry. He was therefore motivated to discover the key equation

$$dR_v(\varphi)/d\varphi = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

$$= \mathbf{i}\,R_v(\varphi)\,.$$

This is *Lie's differential equation*.

Integrating between $\varphi = 0$ and $\varphi = \varphi$, we obtain

$$\int_{\mathbf{I}}^{R_v(\varphi)} dR_v(\varphi)/R_v(\varphi) = \mathbf{i}\int_0^{\varphi} d\varphi$$

so that

$$\ln(R_v(\varphi)/\mathbf{I}) = \mathbf{i}\varphi,$$

or

$$R_v(\varphi) = \mathbf{I}e^{\mathbf{i}\varphi}\text{, the solution of Lie's equation.}$$

Previously, we obtained

$$R_v(\varphi) = \mathbf{I}\cos\varphi + \mathbf{i}\sin\varphi.$$

We have, therefore

$$\mathbf{I}e^{\mathbf{i}\varphi} = \mathbf{I}\cos\varphi + \mathbf{i}\sin\varphi\,.$$

This is an independent proof of the famous Cotes-Euler equation.

We introduce an operator of the form

$$\mathbf{O} = g(x, y, \partial/\partial x, \partial/\partial y),$$

and ask the question: does

$$\delta x = \mathbf{O}f(x, y; \delta\varphi)\ ?$$

Lie answered the question in the affirmative; he found

$$\delta x = \mathbf{O}(x\delta\varphi) = (x\partial/\partial y - y\partial/\partial x)x\delta\varphi = -y\delta\varphi$$

and

$$\delta y = \mathbf{O}(y\delta\varphi) = (x\partial/\partial y - y\partial/\partial x)y\partial\varphi = \ x\delta\varphi\,.$$

Putting $x = x_1$ and $y = x_2$, we obtain

$$\delta x_i = \mathbf{X}x_i\delta\varphi,\ i = 1, 2$$

where

$\mathbf{X} = \mathbf{O} = (x_1\partial/\partial x_2 - x_2\partial/\partial x_1)$, the "generator of rotations" in the plane.

6.3 Exponentiation of infinitesimal rotations

We have seen that

$$R_v(\varphi) = e^{\mathbf{i}\varphi},$$

and therefore

$$R_v(\delta\varphi) = \mathbf{I} + \mathbf{i}\delta\varphi, \text{ for an infinitesimal rotation, } \delta\varphi$$

Performing two infinitesimal rotations in succession, we have

$$R_v^2(\delta\varphi) = (\mathbf{I} + \mathbf{i}\delta\varphi)^2$$

$$= \mathbf{I} + 2\mathbf{i}\delta\varphi \text{ to first order,}$$

$$= R_v(2\delta\varphi).$$

Applying $R_v(\delta\varphi)$ n-times gives

$$R_v^n(\delta\varphi) = R_v(n\delta\varphi) = e^{\mathbf{i}n\delta\varphi} = e^{\mathbf{i}\varphi}$$

$$= R_v(\varphi) \text{ (as } n \to \infty \text{ and } \delta\varphi \to 0, \text{ the}$$

$$\text{product } n\delta\varphi \to \varphi).$$

This result agrees, as it should, with the exact solution of Lie's differential equation.

A finite rotation can be built up by exponentiation of infinitesimal rotations, each one being close to the identity. In general, this approach has the advantage that the infinitesimal form of a transformation can often be found in a straightforward way, whereas the finite form is often intractable.

6.4 Infinitesimal rotations and angular momentum operators

In Classical Mechanics, the angular momentum of a mass m, moving in the plane about the origin of a cartesian reference frame with a momentum **p** is

$$\mathbf{L}_z = \mathbf{r} \times \mathbf{p} = rp\sin\varphi\mathbf{n}_z$$

where $\mathbf{n}_z$ is a unit vector normal to the plane, and φ is the angle between $\mathbf{r}$ and $\mathbf{p}$. In component form, we have

$L_z^{cl} = xp_y - yp_x$, where p_x and p_y are the cartesian components of $\mathbf{p}$.

The transition between Classical and Quantum Mechanics is made by replacing

p_x by $-i(h/2\pi)\partial/\partial x$ (a differential operator)

and p_y by $-i(h/2\pi)\partial/\partial y$ (a differential operator),where h is Planck's constant.

We can therefore write the quantum operator as

$$\mathbf{L}_z^{Q} = -i(h/2\pi)(x\partial/\partial y - y\partial/\partial x) = -i(h/2\pi)\mathbf{X}$$

and therefore

$$\mathbf{X} = i\mathbf{L}_z^{Q}/(h/2\pi),$$

and

$$\delta x_i = \mathbf{X}x_i\,\delta\varphi = (2\pi i\mathbf{L}_z^{Q}/h)x_i\,\delta\varphi, \; i = 1,2.$$

Let an arbitrary, continuous, differentiable function f(x, y) be transformed under the infinitesimal changes

$$x' = x - y\delta\varphi$$

$$y' = y + x\delta\varphi.$$

Using Taylor's theorem, we can write

$$f(x', y') = f(x + \delta x, y + \delta y)$$

$$= f(x - y\delta\varphi, y + x\delta\varphi)$$

$$= f(x, y) + ((\partial f/\partial x)\delta x + ((\partial f/\partial y)\delta y)$$

$$= f(x, y) + \delta\varphi(-y(\partial/\partial x) + x(\partial/\partial y))f(x, y)$$

$$= I + 2\pi i\delta\varphi L_z/h)f(x, y)$$

$$= e^{2\pi i\delta\varphi Lz/h} f(x, y)$$

$$= R_v(2\pi L_z\delta\varphi/h) f(x, y).$$

The invatriance of length under rotations follows at once from this result:

If $f(x, y) = x^2 + y^2$ then

$\partial f/\partial x = 2x$ and $\partial f/\partial y = 2y$, and therefore

$$f(x', y') = f(x, y) + 2x\delta x + 2y\delta y$$

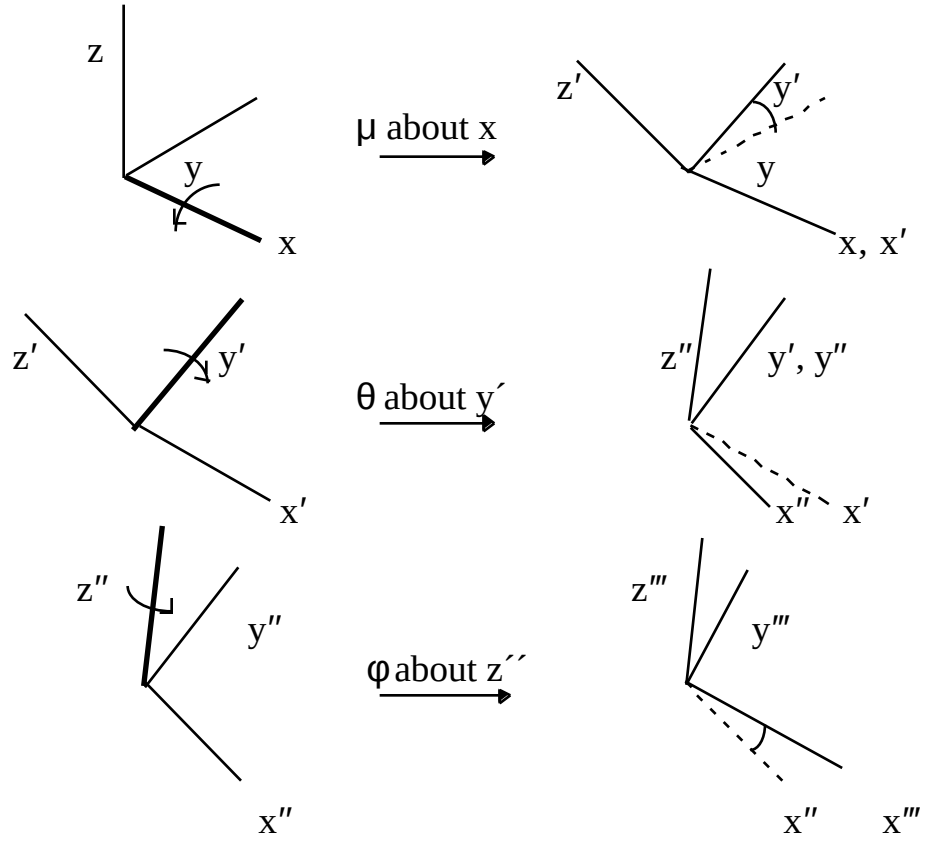
$$= f(x, y) - 2x(y\delta\varphi) + 2y(x\delta\varphi)$$

$$= f(x, y) = x^2 + y^2 = \text{invariant.}$$

This is the only form that leads to the invariance of length under rotations.

6.5 3-dimensional rotations

Consider three successive counterclockwise rotations about the x, y´, and z´´ axes through angles μ, θ, and φ, respectively:



The total transformation is

$$R_c(\mu, \theta, \varphi) = R_c(\varphi) R_c(\theta) R_c(\mu)$$

$$= \begin{pmatrix} \cos\varphi\cos\theta & \cos\varphi\sin\theta\sin\mu + \sin\varphi\cos\mu & -\cos\varphi\sin\theta\cos\mu + \sin\varphi\sin\mu \\ -\sin\varphi\cos\theta & -\sin\varphi\sin\theta\sin\mu + \cos\varphi\cos\mu & \sin\varphi\sin\theta\cos\mu + \sin\varphi\sin\mu \\ \sin\theta & -\cos\theta\sin\mu & \cos\theta\cos\mu \end{pmatrix}$$

For infinitesimal rotations, the total rotation matrix is, to 1st-order in the δ's:

$$R_c(\delta\mu, \delta\theta, \delta\varphi) = \begin{pmatrix} 1 & \delta\varphi & -\delta\theta \\ -\delta\varphi & 1 & \delta\mu \\ \delta\theta & -\delta\mu & 1 \end{pmatrix}.$$

The infinitesimal form can be written as follows:

$$R_c(\delta\mu, \delta\theta, \delta\varphi) = \begin{pmatrix} 1 & \delta\varphi & 0 \\ -\delta\varphi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -\delta\theta \\ 0 & 1 & 0 \\ \delta\theta & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \delta\mu \\ 0 & -\delta\mu & 1 \end{pmatrix}$$

$$= (\mathbf{I} + \mathbf{Y}_3\delta\varphi)(\mathbf{I} + \mathbf{Y}_2\delta\theta)(\mathbf{I} + \mathbf{Y}_1\delta\mu)$$

where

$$\mathbf{Y}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \mathbf{Y}_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{Y}_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

To 1st-order in the δ's, we have

$$R_c(\delta\mu, \delta\theta, \delta\varphi) = \mathbf{I} + \mathbf{Y}_1\delta\mu + \mathbf{Y}_2\delta\theta + \mathbf{Y}_3\delta\varphi.$$

6.6 Algebra of the angular momentum operators

The algebraic properties of the **Y**'s are important. For example, we find that their commutators are:

$$[\mathbf{Y}_1, \mathbf{Y}_2] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

$$= -\mathbf{Y}_3 ,$$

$$[\mathbf{Y}_1, \mathbf{Y}_3] = \mathbf{Y}_2 ,$$

and

$$[\mathbf{Y}_2, \mathbf{Y}_3] = -\mathbf{Y}_1 .$$

These relations define the algebra of the **Y**'s. In general, we have

$$[\mathbf{Y}_j, \mathbf{Y}_k] = \pm \mathbf{Y}_l = \varepsilon_{jkl} \mathbf{Y}_l ,$$

where ε_{jkl} is the anti-symmetric Levi-Civita symbol. It is equal to +1 if jkl is an even permutation, −1 if jkl is an odd permutation, and it is equal to zero if two indices are the same.

Motivated by the relationship between $\mathbf{L}_z$ and $\mathbf{X}$ in 2-dimensions, we introduce the operators

$$\mathbf{J}_k = -i(2\pi/h)\mathbf{Y}_k \text{ , } k = 1, 2, 3.$$

Their commutators are obtained from those of the **Y**'s, for example

$$[\mathbf{Y}_1, \mathbf{Y}_2] = -\mathbf{Y}_3 \rightarrow [2\pi i\mathbf{J}_1/h, 2\pi i\mathbf{J}_2/h] = -2\pi i\mathbf{J}_3/h$$

or

$$-[\mathbf{J}_1, \mathbf{J}_2](2\pi/h)^2 = -2\pi i\mathbf{J}_3/h$$

and therefore

$$[\mathbf{J}_1, \mathbf{J}_2] = ih\mathbf{J}_3/2\pi \text{ .}$$

These operators obey the general commutation relation

$$[\mathbf{J}_j, \mathbf{J}_k] = ih\varepsilon_{jkl}\ \mathbf{J}_l/2\pi \text{ .}$$

The angular momentum operators form a "Lie Algebra".

The basic algebraic properties of the angular momentum operators in Quantum Mechanics stem directly from this relation.

Another approach involves the use of the differential operators in 3-dimensions. A point P[x, y, z] transforms under an infinitesimal rotation of the coordinates as follows

$$\mathbf{P}'[x', y', z'] = R_c(\delta\mu, \delta\theta, \delta\varphi]\mathbf{P}[x, y, z]$$

Substituting the infinitesimal form of R_c in this equation gives

$$\delta x = x' - x = \quad y\delta\varphi - z\delta\theta$$

$$\delta y = y' - y = -x\delta\varphi \quad + z\delta\mu$$

$$\delta z = z' - z = \quad x\delta\theta - y\delta\mu \text{ .}$$

Introducing the classical angular momentum operators: L_i^{cl}, we find that these small changes can be written

$$\delta x_i = \sum_{k=1}^{3} \delta\alpha_k \mathbf{X}_k x_i$$

For example, if i = 1

$$\delta x_1 = \delta x = \delta\mu(z\partial/\partial y - y\partial/\partial z)x + \delta\theta(-z\partial/\partial x + x\partial/\partial z)x + \delta\varphi(y\partial/\partial x - x\partial/\partial y)x = -z\delta\theta + y\delta\varphi.$$

Extending Lie's method to three dimensions, the infinitesimal form of the rotation operator is readily shown to be

$$R_c(\delta\mu, \delta\theta, \delta\varphi) = \mathbf{I} + \sum_{i=1}^{3} (\partial R_c/\partial\alpha_i)\Big|_{\text{All } \alpha i\text{'s} = 0} \delta\alpha_i .$$

7

LIE'S CONTINUOUS TRANSFORMATION GROUPS

In the previous chapter, we discussed the properties of infinitesimal rotations in 2- and 3-dimensions, and we found that they are related directly to the angular momentum operators of Quantum Mechanics. Important algebraic properties of the matrix representations of the operators also were introduced. In this chapter, we shall consider the subject in general terms.

Let x_i, i = 1 to n be a set of n variables. They may be considered to be the coordinates of a point in an n-dimensional vector space, V_n. A set of equations involving the x_i's is obtained by the transformations

$$x_i' = f_i(x_1, x_2, \ldots x_n: a_1, a_2, \ldots a_r),\ i = 1 \text{ to } n$$

in which the set $a_1, a_2, \ldots a_r$ contains r-independent parameters. The set T_a, of transformations maps $x \rightarrow x'$. We shall write

$$x' = f(x; a) \text{ or } \mathbf{x}' = T_a\mathbf{x}$$

for the set of functions.

It is assumed that the functions f_i are differentiable with respect to the x's and the a's to any required order. These functions necessarily depend on the essential parameters, a. This means that no two transformations with different numbers of parameters are the same. r is the smallest number required to characterize the transformation, completely.

The set of functions f_i forms a finite continuous group if:

1. The result of two successive transformations $x \rightarrow x' \rightarrow x''$ is equivalent to a single transformation $x \rightarrow x''$:

$$\begin{aligned} x' = f(x'; b) &= f(f(x; a); b) \\ &= f(x; c) \\ &= f(x; \chi(a; b)) \end{aligned}$$

where c is the set of parameters

$$c_\lambda = \chi_\lambda (a; b)\ ,\ \lambda = 1 \text{ to } r,$$

and

2. To every transformation there corresponds a unique inverse that belongs to the set:

$\bar{a}$ such that $x = f(x'; \bar{a}) = \bar{f}(x'; a)$

We have

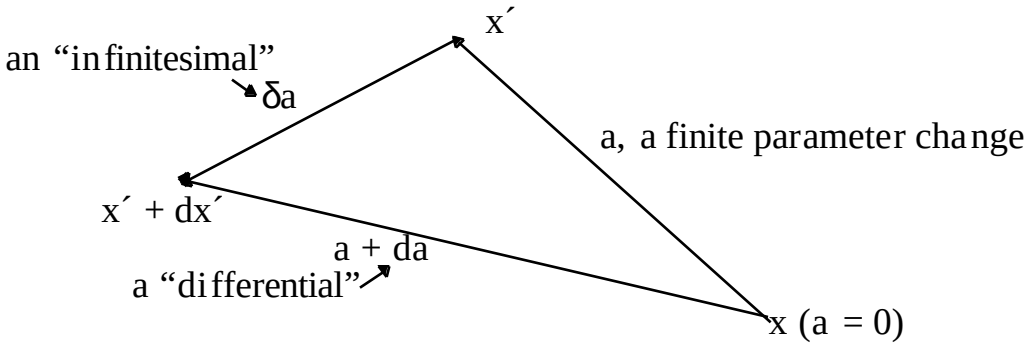
$$T_a T_a^{-1} = T_a^{-1} T_a = I, \text{ the identity.}$$

We shall see that 1) is a highly restrictive requirement.

The transformation $x = f(x; a_0)$ is the identity. Without loss of generality, we can take $a_0 = 0$. The essential point of Lie's theory of continuous transformation groups is to consider that part of the group that is close to the identity, and not to consider the group as a whole. Successive infinitesimal changes can be used to build up the finite change.

7.1 One-parameter groups

Consider the transformation $x \rightarrow x'$ under a finite change in a single parameter a, and then a change $x' + dx'$. There are two paths from $x \rightarrow x' + dx'$; they are as shown:



We have

$$x' + dx' = f(x; a + da)$$

$$= f(f(x; a); \delta a) = f(x'; \delta a)$$

The 1st-order Taylor expansion is

$$dx' = \partial f(x'; a)/\partial a \,\Big|_{a=0} \delta a \equiv u(x')\,\delta a$$

The Lie group conditions then demand

$$a + da = \chi(a; \delta a).$$

But

$$\chi(a; 0) = a, \quad (b = 0)$$

therefore

$$a + da = a + \partial\chi(a; b)/\partial b\,\Big|_{b=0} \delta a$$

so that

$$da = \partial\chi(a; b)/\partial b\,\Big|_{b=0} \delta a$$

or

$$\delta a = A(a)da.$$

Therefore

$$dx' = u(x')A(a)da,$$

leading to

$$dx'/u(x') = A(a)da$$

so that

$$\int_x^{x'} dx'/u(x') = \int_0^a A(a)da \equiv s, \; (s = 0 \rightarrow \text{the identity}).$$

We therefore obtain

$$U(x') - U(x) = s.$$

A transformation of coordinates (new variables) therefore transfers all elements of the group by the same transformation: *a one-parameter group is equivalent to a group of translations.*

Two continuous transformation groups are said to be similar when they can be obtained from one another by a change of variable. For example, consider the group defined by

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

The identity coprresponds to a = 1. The infinitesimal transformation is therefore

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} (1 + \delta a) & 0 \\ 0 & (1 + \delta a)^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

To 1st-order in δa we have

$$x_1' = x_1 + x_1 \delta a$$

and

$$x_2' = x_2 + 2x_2 \delta a$$

or

$$\delta x_1 = x_1 \delta a$$

and

$$\delta x_2 = 2x_2 \delta a.$$

In the limit, these equations give

$$dx_1/x_1 = dx_2/2x_2 = da.$$

These are the differential equations that correspond to the infinitesimal equations above.

Integrating, we have

$$\int_{x1}^{x1'} dx_1/x_1 = \int_0^a da \quad \text{and} \quad \int_{x2}^{x2'} dx_2/2x_2 = \int_0^a da \ ,$$

so that

$$\ln x_1' - \ln x_1 = a = \ln(x_1'/x_1)$$

and

$$\ln(x_2'/x_2) = 2a = 2\ln(x_1'/x_1)$$

or

$$U' = (x_2'/x_1'^2) = U = (x_2/x_1^2) \ .$$

Putting $V = \ln x_1$, we obtain

$$V' = V + a \quad \text{and } U' = U, \text{ the } \textit{translation group}.$$

7.2 Determination of the finite equations from the infinitesimal forms

Let the finite equations of a one-parameter group $G_{(1)}$ be

$$x_1' = \varphi(x_1, x_2 ; a)$$

and

$$x_2' = \psi(x_1, x_2 ; a),$$

and let the identity correspond to $a = 0$.

We consider the transformation of $f(x_1, x_2)$ to $f(x_1', x_2')$. We expand $f(x_1', x_2')$ in a Maclaurin series in the parameter a (at definite values of x_1 and x_2):

$$f(x_1', x_2') = f(0) + f'(0)a + f''(0)a^2/2! + \ ...$$

where

$$f(0) = f(x_1', x_2')|_{a=0} = f(x_1, x_2),$$

and

$$f'(0) = (df(x_1', x_2')/da|_{a=0}$$
$$= \{(\partial f/\partial x_1')(dx_1'/da) + (\partial f/\partial x_2')(dx_2'/da)\}|_{a=0}$$
$$= \{(\partial f/\partial x_1')u(x_1', x_2') + (\partial f/\partial x_2')v(x_1', x_2')\}|_{a=0}$$

therefore

$$f'(0) = \{(u(\partial/\partial x_1) + v(\partial/\partial x_2))f\}|_{a=0}$$
$$= \mathbf{X}f(x_1, x_2).$$

Continuing in this way, we have

$$f''(0) = \{d^2f(x_1', x_2')/da^2\}|_{a=0} = \mathbf{X}^2f(x_1, x_2), \text{ etc....}$$

The function $f(x_1', x_2')$ can be expanded in the series

$$f(x_1', x_2') = f(0) + af'(0) + (a^2/2!)f''(0) + ...$$
$$= f(x_1, x_2) + a\mathbf{X}f + (a^2/2!)\mathbf{X}^2f + ...$$

$\mathbf{X}^n$f is the symbol for operating n-times in succession of f with **X**.

The finite equations of the group are therefore

$$x_1' = x_1 + a\mathbf{X}x_1 + (a^2/2!)\mathbf{X}^2x_1 + ...$$

and

$$x_2' = x_2 + a\mathbf{X}x_2 + (a^2/2!)\mathbf{X}^2x_2 + = ...$$

If x_1 and x_2 are definite values to which $x_1'$and x_2' reduce for the identity a=0, then these equations are the series solutions of the differential equations

$$dx_1'/u(x_1', x_2') = dx_2'/v(x_1', x_2') = da.$$

*The group is referred to as the group **Xf**.*

For example, let

$$\mathbf{X}f = (x_1\partial/\partial x_1 + x_2\partial/\partial x_2)f$$

then

$$\begin{aligned} x_1' &= x_1 + a\mathbf{X}x_1 + (a^2/2!)\mathbf{X}^2 f \ldots \\ &= x_1 + a(x_1\partial/\partial x_1 + x_2\partial/\partial x_2)x_1 + \ldots \\ &= x_1 + ax_1 + (a^2/2!)(x_1\partial/\partial x_1 + x_2\partial/\partial x_2)x_1 + \\ &= x_1 + ax_1 + (a^2/2!)x_1 + \ldots \\ &= x_1(1 + a + a^2/2! + \ldots) \\ &= x_1 e^a. \end{aligned}$$

Also, we find

$$x_2' = x_2 e^a.$$

Putting $b = e^a$, we have

$$x_1' = bx_1, \text{ and } x_2' = bx_2.$$

The finite group is the group of magnifications.

If $\mathbf{X} = (x\partial/\partial y - y\partial/\partial x)$ we find, for example, that the finite group is the group of 2-dimensional rotations.

7.3 Invariant functions of a group

Let

$\mathbf{X}f = (u\partial/\partial x_1 + v\partial/\partial x_2)f$ define a one-parameter group, and let a=0 give the identity. A function $F(x_1, x_2)$ is termed an *invariant* under the transformation group $G_{(1)}$ if

$$F(x_1', x_2') = F(x_1, x_2)$$

for all values of the parameter, a.

The function $F(x_1', x_2')$ can be expanded as a series in a:

$$F(x_1', x_2') = F(x_1, x_2) + a\mathbf{X}F + (a^2/2!)\mathbf{X}(\mathbf{X}F) + \ldots$$

If

$$F(x_1', x_2') = F(x_1, x_2) = \text{invariant for all values of a,}$$

it is necessary for

$$\mathbf{X}F = 0,$$

and this means that

$$\{u(x_1, x_2)\partial/\partial x_1 + v(x_1, x_2)\partial/\partial x_2\}F = 0.$$

Consequently,

$$F(x_1, x_2) = \text{constant}$$

is a solution of

$$dx_1/u(x_1, x_2) = dx_2/v(x_1, x_2) .$$

This equation has one solution that depends on one arbitrary constant, and therefore $G_{(1)}$ has only one basic invariant, and all other possible invariants can be given in terms of the basic invariant.

For example, we now reconsider the the invariants of rotations:

The infinitesimal transformations are given by

$$\mathbf{X}f = (x_1\partial/\partial x_2 - x_2\partial/\partial x_1),$$

and the differential equation that gives the invariant function F of the group is obtained by solving the characteristic differential equations

$$dx_1/x_2 = d\varphi, \text{ and } dx_2/x_1 = -d\varphi,$$

so that

$$dx_1/x_2 + dx_2/x_1 = 0.$$

The solution of this equation is

$$x_1^2 + x_2^2 = \text{constant},$$

and therefore the invariant function is

$$F(x_1, x_2) = x_1^2 + x_2^2.$$

All functions of $x_1^2 + x_2^2$ are therefore invariants of the 2-dimensional rotation group.

This method can be generalized. A group $G_{(1)}$ in n-variables defined by the equation

$$x_i' = \varphi(x_1, x_2, x_3, \ldots x_n; a),\ i = 1 \text{ to } n,$$

is equivalent to a unique infinitesimal transformation

$$\mathbf{X}f = u_1(x_1, x_2, x_3, \ldots x_n)\partial f/\partial x_1 + \ldots u_n(x_1, x_2, x_3, \ldots x_n)\partial f/\partial x_n .$$

If a is the group parameter then the infinitesimal transformation is

$$x_i' = x_i + u_i(x_1, x_2, \ldots x_n)\delta a \quad (i = 1 \text{ to } n),$$

then, if $E(x_1, x_2, \ldots x_n)$ is a function that can be differentiated n-times with respect to its arguments, we have

$$E(x_1', x_2', \ldots x_n') = E(x_1, x_2, \ldots x_n) + a\mathbf{X}E + (a^2/2!)\mathbf{X}^2E + .$$

Let $(x_1, x_2, \ldots x_n)$ be the coordinates of a point in n-space and let a be a parameter, independent of the x_i's. As a varies, the point $(x_1, x_2, \ldots x_n)$ will describe a trajectory, starting from the initial point $(x_1, x_2, \ldots x_n)$. A necessary and sufficient condition that $F(x_1, x_2, \ldots x_n)$ be an invariant function is that $\mathbf{X}F = 0$. A curve $F = 0$ is a trajectory and therefore an invariant curve if

$$\mathbf{X}F(x_1, x_2, x_3, \ldots x_n) = 0.$$

8

PROPERTIES OF n-VARIABLE, r-PARAMETER LIE GROUPS

The change of an n-variable function F(**x**) produced by the infinitesimal transformations associated with r-essential parameters is:

$$dF = \sum_{i=1}^{n} (\partial F/\partial x_i) dx_i$$

where

$$dx_i = \sum_{\lambda=1}^{r} u_{i\lambda}(\mathbf{x})\delta a_\lambda \text{ , the Lie form.}$$

The parameters are independent of the x_i's therefore we can write

$$dF = \sum_{\lambda=1}^{r} \delta a_\lambda \{\sum_{i=1}^{n} u_{i\lambda}(\mathbf{x})(\partial/\partial x_i)F\}$$

$$= \sum_{\lambda=1}^{r} \delta a_\lambda \, \mathbf{X}_\lambda F$$

where the infinitesimal generators of the group are

$$X_\lambda \equiv \sum_{i=1}^{n} u_{i\lambda}(\mathbf{x})(\partial/\partial x_i) \text{ , } \lambda = 1 \text{ to r.}$$

The operator

$$\mathbf{I} + \sum_{\lambda=1}^{r} \mathbf{X}_\lambda \delta a_\lambda$$

differs infinitesimally from the identity.

The generators $\mathbf{X}_\lambda$ have algebraic properties of basic importance in the Theory of Lie Groups. The $\mathbf{X}_\lambda$'s are differential operators. The problem is therefore one of obtaining the algebraic structure of differential operators. This problem has its origin in the work of Poisson (1807); he introduced the following ideas:

The two expressions

$$\mathbf{X}_1 f = (u_{11}\partial/\partial x_1 + u_{12}\partial/\partial x_2)f$$

and

$$\mathbf{X}_2 f = (u_{21}\partial/\partial x_1 + u_{22}\partial/\partial x_2)f$$

where the coefficients $u_{i\lambda}$ are functions of the variables x_1, x_2, and $f(x_1, x_2)$ is an arbitrary differentiable function of the two variables, are termed linear differential operators.

The "product" in the order $\mathbf{X}_2$ followed by $\mathbf{X}_1$ is defined as

$$\mathbf{X}_1\mathbf{X}_2 f = (u_{11}\partial/\partial x_1 + u_{12}\partial/\partial x_2)(u_{21}\partial f/\partial x_1 + u_{22}\partial f/\partial x_2)$$

The product in the reverse order is defined as

$$\mathbf{X}_2\mathbf{X}_1 f = (u_{21}\partial/\partial x_1 + u_{22}\partial/\partial x_2)(u_{11}\partial f/\partial x_1 + u_{12}\partial f/\partial x_2).$$

The difference is

$$\begin{aligned}\mathbf{X}_1\mathbf{X}_2 f - \mathbf{X}_2\mathbf{X}_1 f = {} & \mathbf{X}_1 u_{21}\partial f/\partial x_1 + \mathbf{X}_1 u_{22}\partial f/\partial x_2 \\ & - \mathbf{X}_2 u_{11}\partial f/\partial x_1 - \mathbf{X}_2 u_{12}\partial f/\partial x_2. \\ = {} & (\mathbf{X}_1 u_{21} - \mathbf{X}_2 u_{11})\partial f/\partial x_1 + (\mathbf{X}_1 u_{22} - \mathbf{X}_2 u_{12})\partial f/\partial x_2 \\ \equiv {} & [\mathbf{X}_1, \mathbf{X}_2]f.\end{aligned}$$

This quantity is called the *Poisson operator* or the *commutator* of the operators $\mathbf{X}_1 f$ and $\mathbf{X}_2 f$.

The method can be generalized to include $\lambda = 1$ to r essential parameters and i = 1 to n variables. The ath-linear operator is then

$$\begin{aligned}\mathbf{X}_a &= u_{ia}\partial f/\partial x_i \\ &= \sum_{i=1}^{n} u_{ia}\partial f/\partial x_i \text{ , (a sum over repeated indices).}\end{aligned}$$

Lie's differential equations have the form

$$\partial x_i/\partial a_\lambda = u_{ik}(x)A_{k\lambda}(a) \text{ , } i = 1 \text{ to } n, \lambda = 1 \text{ to } r.$$

Lie showed that

$$(\partial c_{k\tau\sigma}/\partial a_\rho)u_{ik} = 0$$

in which

$$u_{j\sigma}\partial u_{i\tau}/\partial x_j - u_{j\tau}\partial u_{i\sigma}/\partial x_j = c_{k\tau\sigma}(a)u_{ik}(x),$$

so that the $c_{k\tau\sigma}$'s are *constants.* Furthermore, the commutators can be written

$$[\mathbf{X}_\rho, \mathbf{X}_\sigma] = (c_{k\rho\sigma}u_{jk})\partial/\partial x_j$$

$$= c_{k\rho\sigma}\mathbf{X}_k.$$

The commutators are linear combinations of the $\mathbf{X}_k$'s. (Recall the earlier discussion of the angular momentum operators and their commutators). The $c_{k\rho\sigma}$'s are called the *structure constants* of the group. They have the properties

$$c_{k\rho\sigma} = -c_{k\sigma\rho},$$

$$c_{\mu\rho\sigma}c_{\nu\mu\tau} + c_{\mu\sigma\tau}c_{\nu\mu\rho} + c_{\mu\tau\rho}c_{\nu\mu\sigma} = 0.$$

Lie made the remarkable discovery that, given these structure constants, the functions that satisfy

$$\partial x_i/\partial a_\lambda = u_{ik}A_{k\lambda}(a) \text{ can be found.}$$

(Proofs of all the above important statements, together with proofs of Lie's three fundamental theorems, are given in Eisenhart's standard work *Continuous Groups of Transformations,* Dover Publications, 1961).

8.1 The rank of a group

Let **A** be an operator that is a linear combination of the generators of a group, $\mathbf{X}_i$:

$$\mathbf{A} = \alpha_i\mathbf{X}_i \text{ (sum over i),}$$

and let

$$\mathbf{X} = x_j \mathbf{X}_j .$$

The rank of the group is defined as the minimum number of commuting, linearly independent operators of the form **A**.

We therefore require all solutions of

$$[\mathbf{A}, \mathbf{X}] = 0.$$

For example, consider the orthogonal group, $O^+(3)$; here

$$\mathbf{A} = \alpha_i \mathbf{X}_i \quad i = 1 \text{ to } 3,$$

and

$$\mathbf{X} = x_j \mathbf{X}_j \quad j = 1 \text{ to } 3$$

so that

$$[\mathbf{A}, \mathbf{X}] = \alpha_i x_j [\mathbf{X}_i, \mathbf{X}_j] \quad i, j = 1 \text{ to } 3$$

$$= \alpha_i x_j \varepsilon_{ijk} \mathbf{X}_k .$$

The elements of the sets of generators are linearly independent, therefore

$$\alpha_i x_j \varepsilon_{ijk} = 0 \text{ (sum over i, j,, k = 1, 2, 3)}$$

This equation represents the equations

$$\begin{pmatrix} -\alpha_2 & \alpha_1 & 0 \\ \alpha_3 & 0 & -\alpha_2 \\ 0 & -\alpha_3 & \alpha_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The determinant of α is zero, therefore a non-trivial solution of the x_j's exists. The solution is given by

$$x_j = \alpha_j \quad (j = 1, 2, 3)$$

so that

$$\mathbf{A} = \mathbf{X} .$$

$O^+(3)$ is a group of rank one.

8.2 The Casimir operator of $O^+(3)$

The generators of the rotation group $O^+(3)$ are the operators. $\mathbf{Y}_k$'s, discussed previously. They are directly related to the angular momentum operators, $\mathbf{J}_k$:

$$\mathbf{J}_k = -i(h/2\pi)\mathbf{Y}_k \ (k = 1, 2, 3).$$

The matrix representations of the $\mathbf{Y}_k$'s are

$$\mathbf{Y}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \mathbf{Y}_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{Y}_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The square of the total angular momentum, J is

$$\begin{aligned} \mathbf{J}^2 &= \sum_1^3 \mathbf{J}_i^2 \\ &= (h/2\pi)^2 (\mathbf{Y}_1^2 + \mathbf{Y}_2^2 + \mathbf{Y}_3^2) \\ &= (h/2\pi)^2(-2\mathbf{I}). \end{aligned}$$

Schur's lemma states that an operator that is a constant multiple of $\mathbf{I}$ commutes with all matrix irreps of a group, so that

$$[\mathbf{J}_k, \mathbf{J}^2] = 0 \ , \ k = 1,2 \ ,3.$$

The operator $\mathbf{J}^2$ with this property is called the Casimir operator of the group $O^+(3)$.

In general, the set of operators $\{\mathbf{C}_i\}$ in which the elements commute with the elements of the set of irreps of a given group, forms the set of Casimir operators of the group. All Casimir operators are constant multiples of the unit matrix:

$\mathbf{C}_i = a_i\mathbf{I}$; the constants a_i are characteristic of a particular representation of a group.

9

MATRIX REPRESENTATIONS OF GROUPS

Matrix representations of linear operators are important in Linear Algebra; we shall see that they are equally important in Group Theory.

If a group of m × m matrices

$$\mathbf{D}_n^{(m)} = \{\mathbf{D}_1^{(m)}(g_1),...\mathbf{D}_k^{(m)}(g_k), ...\mathbf{D}_n^{(m)}(g_n)\}$$

can be found in which each element is associated with the corresponding element g_k of a group of order n

$$G_n = \{g_1,...g_k,....g_n\},$$

and the matrices obey

$$\mathbf{D}_j^{(m)}(g_j)\mathbf{D}_i^{(m)}(g_i) = \mathbf{D}_{ji}^{(m)}(g_j g_i),$$

and

$$\mathbf{D}_1^{(m)}(g_1) = \mathbf{I}, \text{ the identity,}$$

then the matrices $\mathbf{D}_k^{(m)}(g_k)$ are said to form an m-dimensional representation of G_n. If the association is one-to-one we have an isomorphism and the representation is said to be *faithful*.

The subject of Group Representations forms a very large branch of Group Theory. There are many standard works on this topic (see the bibliography), each one containing numerous definitions, lemmas and theorems. Here, a rather brief account is given of some of the more important results. The reader should delve into the deeper aspects of the subject as the need arises. The subject will be introduced by considering

representations of the rotation groups, and their corresponding cyclic groups.

9.1 The 3-dimensional representation of rotations in the plane

The rotation of a vector through an angle φ in the plane is characterized by the 2 x 2 matrix

$$R_v(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

The group of symmetry transformations that leaves an equilateral triangle invariant under rotations in the plane is of order three, and each element of the group is of dimension two

$$G_n \sim R_3^{(2)} = \{R(0),\ R(2\pi/3),\ R(4\pi/3)\}$$

$$= \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}, \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{pmatrix} \right\}$$

$$\approx \{123, 312, 231\} = C_3.$$

These matrices form a 2-dimensional representation of C_3 .

A 3-dimensional representation of C_3 can be obtained as follows:

Consider an equilateral triangle located in the plane and let the coordinates of the three vertices $\mathbf{P}_1[x, y]$, $\mathbf{P}_2[x', y']$, and $\mathbf{P}_3[x'', y'']$ be written as a 3-vector $\mathbf{P}_{13} = [\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3]$, in normal order. We introduce 3×3 matrix operators $\mathbf{D}_i^{(3)}$ that change the order of the elements of $\mathbf{P}_{13}$, cyclically. The identity is

$$\mathbf{P}_{13} = \mathbf{D}_1^{(3)}\mathbf{P}_{13}, \text{ where } \mathbf{D}_1^{(3)} = \text{diag}(1, 1, 1).$$

The rearrangement

$$\mathbf{P}_{13} \rightarrow \mathbf{P}_{23}[\mathbf{P}_3, \mathbf{P}_1, \mathbf{P}_2] \text{ is given by}$$

$$\mathbf{P}_{23} = \mathbf{D}_2^{(3)}\mathbf{P}_{13},$$

where

$$\mathbf{D}_2^{(3)} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

and the rearrangement

$$\mathbf{P}_{13} \rightarrow \mathbf{P}_{33}[\mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_1] \text{ is given by}$$

$$\mathbf{P}_{33} = \mathbf{D}_3^{(3)}\mathbf{P}_{13}$$

where

$$\mathbf{D}_3^{(3)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

The set of matrices $\{\mathbf{D}_i^{(3)}\} = \{\mathbf{D}_1^{(3)}, \mathbf{D}_2^{(3)}, \mathbf{D}_3^{(3)}\}$ is said to form a *3-dimensional representation* of the original 2-dimensional representation $\{R_3^{(2)}\}$. The elements $\mathbf{D}_i^{(3)}$ have the same group multiplication table as that associated with C_3.

9.2 The m-dimensional representation of symmetry transformations in d-dimensions

Consider the case in which a group of order n

$$G_n = \{g_1, g_2, \ldots g_k, \ldots g_n\}$$

is represented by

$$R_n^{(m)} = \{ R_1^{(m)}, R_2^{(m)}, ...R_n^{(m)}$$

where

$$R_n^{(m)} \sim G_n,$$

and $R_k^{(m)}$ is an m × m matrix representation of g_k. Let $\mathbf{P}_{1d}$ be a vector in d-dimensional space, written in normal order:

$$\mathbf{P}_{1d} = [\mathbf{P}_1, \mathbf{P}_2, ...\mathbf{P}_d],$$

and let

$$P_{1m} = [\mathbf{P}_{1d}, \mathbf{P}_{2d}, ...\mathbf{P}_{md}]$$

be an m-vector, written in normal order, in which the components are each d-vectors. Introduce the m × m matrix operator $\mathbf{D}_k^{(m)}(g_k)$ such that

$$P_{1m} = \mathbf{D}_1^{(m)}(g_1) P_{1m}$$

$$P_{2m} = \mathbf{D}_2^{(m)}(g_2) P_{1m}$$

.

.

$P_{km} = D_k^{(m)}(g_k) P_{1m}$, k = 1 to m, the number of symmetry operations,

where P_{km} is the kth (cyclic) permutation of P_{1m} , and $\mathbf{D}_k^{(m)}(g_k)$ is called the "m-dimensional representation of g_k".

Infinitely many representations of a given representation can be found, for, if **S** is a matrix representation, and **M** is any definite matrix with an inverse, we can form $\mathbf{T}(x) = \mathbf{M}\mathbf{S}(x)\mathbf{M}^{-1}$, $x \in G$. Since

$$\mathbf{T}(xy) = \mathbf{M}\mathbf{S}(xy)\mathbf{M}^{-1} = \mathbf{M}\mathbf{S}(x)\mathbf{S}(y)\mathbf{M}^{-1} = \mathbf{M}\mathbf{S}(x)\mathbf{M}^{-1}\mathbf{M}\mathbf{S}(y)\mathbf{M}^{-1}$$

$$= \mathbf{T}(x)\mathbf{T}(y),$$

T is a representation of G. The new representation simply involves a change of variable in the corresponding substitutions. Representations related in the manner of **S** and **T** are *equivalent* , and are not regarded as different representations. All representations that are equivalent to **S** are equivalent to each other, and they form an infinite class. Two equivalent representations will be written $\mathbf{S} \sim \mathbf{T}$.

9.3 Direct sums

If **S** is a representation of dimension s, and **T** is a representation of dimension t of a group G, the matrix

$$\mathbf{P} = \begin{pmatrix} \mathbf{S}(g) & \mathbf{0} \\ \mathbf{0} & \mathbf{T}(g) \end{pmatrix}, \quad (g \in G)$$

of dimension s + t is called the *direct sum* of the matrices **S**(g) and **T**(g), written $\mathbf{P} = \mathbf{S} \oplus \mathbf{T}$. Therefore, given two representations (they can be the same), we can obtain a third by adding them directly. Alternatively, let **P** be a representation of dimension s + t; we suppose that, for all $x \in G$, the matrix **P**(x) is of the form

$$\begin{pmatrix} \mathbf{A}(x) & \mathbf{0} \\ \mathbf{0} & \mathbf{B}(x) \end{pmatrix}$$

where $\mathbf{A}(x)$ and $\mathbf{B}(x)$ are $s \times s$ and $t \times t$ matrices, respectively. (The $\mathbf{0}$'s are $s \times t$ and $t \times s$ zero matrices). Define the matrices $\mathbf{S}$ and $\mathbf{T}$ as follows:

$$\mathbf{S}(x) \equiv \mathbf{A}(x) \text{ and } \mathbf{T}(x) \equiv \mathbf{B}(x), \quad x \quad G.$$

Since, by the group property, $\mathbf{P}(xy) = \mathbf{P}(x)\mathbf{P}(y)$,

$$\begin{pmatrix} \mathbf{A}(xy) & \mathbf{0} \\ \mathbf{0} & \mathbf{B}(xy) \end{pmatrix} = \begin{pmatrix} \mathbf{A}(x) & \mathbf{0} \\ \mathbf{0} & \mathbf{B}(x) \end{pmatrix} \begin{pmatrix} \mathbf{A}(y) & \mathbf{0} \\ \mathbf{0} & \mathbf{B}(y) \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{A}(x)\mathbf{A}(y) & \mathbf{0} \\ \mathbf{0} & \mathbf{B}(x)\mathbf{B}(y) \end{pmatrix}.$$

Therefore, $\mathbf{S}(xy) = \mathbf{S}(x)\mathbf{S}(y)$ and $\mathbf{T}(xy) = \mathbf{T}(x)\mathbf{T}(y)$, so that $\mathbf{S}$ and $\mathbf{T}$ are representations. The representation $\mathbf{P}$ is said to be *decomposable*, with components $\mathbf{S}$ and $\mathbf{T}$. A representation is indecomposable if it cannot be decomposed.

If a component of a decomposable representation is itself decomposable, we can continue in this manner to decompose any representation into a finite number of indecomposable components. (It should be noted that the property of indecomposablity depends on the field of the representation; the real field must sometimes be extended to the complex field to check for indecomposability).

A weaker form of decomposability arises when we consider a matrix of the form

$$\mathbf{P}(x) = \begin{pmatrix} \mathbf{A}(x) & \mathbf{0} \\ \mathbf{E}(x) & \mathbf{B}(x) \end{pmatrix}$$

where **A**(x), and **B**(x) are matrices of dimensions s × s and t × t respectively and **E**(x) is a matrix that depends on x, and **0** is the s × t zero matrix. The matrix **P**, and any equivalent form, is said to be *reducible*. An *irreducible* representation is one that cannot be reduced. Every decomposable matrix is reducible (**E**(x) = **0**), whereas a reducible representation need not be decomposable.

If **S** and **T** are reducible, we can continue in this way to obtain a set of irreducible components. The components are determined uniquely, up to an equivalence. The set of distinct irreducible representations of a finite group is (in a given field) an *invariant* of the group. The components form the *building blocks* of a representation of a group.

In Physics, *decomposable* representations are generally referred to as *reducible representations* (reps).

9.4 Similarity and unitary transformations and matrix diagonalization

Before discussing the question of the possibility of reducing the dimension of a given representation, it will be useful to consider some important results in theTheory of Matrices. The proofs of these statements are given in the standard works on Matrix Theory. (See bibliography).

If there exists a matrix **Q** such that

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{B},$$

then the matrices **A** and **B** are related by a *similarity transformation*.

If **Q** is unitary ($\mathbf{QQ}^{\dagger} = \mathbf{I}$: $\mathbf{Q}^{\dagger} = (\mathbf{Q}^*)^T$, the hermitian conjugate) then **A** and **B** are related by a *unitary transformation*.

If $\mathbf{A}' = \mathbf{Q}^{-1}\mathbf{AQ}$; $\mathbf{B}' = \mathbf{Q}^{-1}\mathbf{BQ}$; $\mathbf{C}' = \mathbf{Q}^{-1}\mathbf{CQ}$..then any algebraic relation among **A**, **B**, **C**...is also satisfied by **A**′, **B**′, **C**′ ...

If a similarity transformation produces a diagonal matrix then the process is called *diagonalization*.

If **A** and **B** can be diagonalized by the same matrix then **A** and **B** *commute*.

If **V** is formed from the eigenvectors of **A** then the similarity transformation $\mathbf{V}^{-1}\mathbf{AV}$ will produce a diagonal matrix whose elements are the eigenvalues of **A**.

If **A** is hermitian then **V** will be unitary and therefore an hermitian matrix can always be diagonalized by a unitary transformation. A real symmetric matrix can always be diagonalized by an orthogonal transformation.

9.5 The Schur-Auerbach theorem

This theorem states

Every matrix representation of a finite group is equivalent to a unitary matrix representation

Let $G_n = \{\mathbf{D}_1, \mathbf{D}_2, ...\mathbf{D}_n\}$ be a matrix group, and let **D** be the matrix formed by taking the sum of pairs of elements

$$\mathbf{D} = \sum_{i=1}^{n} \mathbf{D}_i \mathbf{D}_i^{\dagger}$$

where $\mathbf{D}_i^{\dagger}$ is the hermitian conjugate of $\mathbf{D}_i$.

Since $\mathbf{D}_i$ is non-singular, each term in the sum is positive definite. Therefore **D** itself is positive definite. Let $\mathbf{L}_d$ be a diagonal matrix that is equivalent to **D**, and let $\mathbf{L}_d^{1/2}$ be the positive definite matrix formed by replacing the elements of $\mathbf{L}_d$ by their positive square roots. Let **U** be a unitary matrix with the property that

$$\mathbf{L}_d = \mathbf{U}\mathbf{D}\mathbf{U}^{-1}.$$

Introduce the matrix

$$\mathbf{S} = \mathbf{L}_d^{-1/2}\mathbf{U},$$

then $\mathbf{S}\mathbf{D}_i\mathbf{S}^{-1}$ is *unitary*. (This property can be demonstrated by considering $(\mathbf{S}\mathbf{D}_i\mathbf{S}^{-1})(\mathbf{S}\mathbf{D}_i\mathbf{S}^{-1})^{\dagger}$, and showing that it is equal to the identity.). **S** will transform the original matrix representation G_n into diagonal form. Every unitary matrix is diagonalizable, and therefore every matrix in every finite matrix representation can be diagonalized.

9.6 Schur's lemmas

A matrix representation is reducible if every element of the representation can be put in block-diagonal form by a single similarity transformation. Invoking the result of the previous section, we need only discuss unitary representations.

If $G_n = \{\mathbf{D}^{(\nu)}(R)\}$ is an *irreducible* representation of dimension ν of a group G_n, and $\{\mathbf{D}^{(\mu)}(R)\}$ is an *irreducible* representation of dimension μ of the same group, G_n, and if there exists a matrix **A** such that

$$\mathbf{D}^{(\nu)}(R)\mathbf{A} = \mathbf{A}\mathbf{D}^{(\mu)}(R) \qquad R \quad G_n$$

then either

i) $\mathbf{A} = \mathbf{0}$

or

ii) $\mathbf{A}$ is a square non-singular matrix (so that $\nu = \mu$)

Let the μ columns of $\mathbf{A}$ be written $\mathbf{c}_1, \mathbf{c}_2, \ldots\mathbf{c}_\mu$, then, for any matrices $\mathbf{D}^{(\nu)}$ and $\mathbf{D}^{(\mu)}$ we have

$$\mathbf{D}^{(\nu)}\mathbf{A} = (\mathbf{D}^{(\nu)}\mathbf{c}_1, \mathbf{D}^{(\nu)}\mathbf{c}_2, \ldots\mathbf{D}^{(\nu)}\mathbf{c}_n)$$

an

$$\mathbf{A}\mathbf{D}^{(\mu)} = \left(\sum_{k=1}^{\mu}\mathbf{D}^{(\mu)}_{k1}\mathbf{c}_k, \sum_{k=1}^{\mu}\mathbf{D}^{(\mu)}_{k2}\mathbf{c}_k, \ldots\sum_{k=1}^{\mu}\mathbf{D}^{(\mu)}_{k\mu}\mathbf{c}_k\right).$$

therefore

$$\mathbf{D}^{(\nu)}\mathbf{c}_j = \sum_{k=1}^{\mu}\mathbf{D}^{(\mu)}_{kj}\mathbf{c}_k$$

and therefore the μ $\mathbf{c}$-vectors span a space that is invariant under the irreducible set of ν-dimensional matrices $\{\mathbf{D}^{(\nu)}\}$. The $\mathbf{c}$-vectors are therefore the null-vector *or* they span a ν-dimensional vector space. The first case corresponds to $\mathbf{A} = \mathbf{0}$, and the second to $\mu \geq \nu$ and $\mathbf{A} \neq \mathbf{0}$.

In the second case, the hermitian conjugates $\mathbf{D}^{(\nu)}{}_1^\dagger, \ldots\mathbf{D}^{(\nu)}{}_n^\dagger$ and $\mathbf{D}^{(\mu)}{}_1^\dagger, \ldots\mathbf{D}^{(\mu)}{}_n^\dagger$ also are *irreducible* . Furthermore, since $\mathbf{D}^{(\nu)}{}_i(R)\mathbf{A} = \mathbf{A}\mathbf{D}^{(\mu)}{}_i(R)$

$$\mathbf{D}^{(\mu)}{}_i^\dagger\mathbf{A}^\dagger = \mathbf{A}^\dagger\mathbf{D}^{(\nu)}{}_i^\dagger,$$

and therefore, following the method above, we find that $\nu \geq \mu$. We must therefore have $\nu = \mu$, so that $\mathbf{A}$ is square.. Since the ν-columns of $\mathbf{A}$ span a ν-dimensional space, the matrix $\mathbf{A}$ is necessarily non-singular. □

As a corollary, a matrix $\mathbf{D}$ that *commute*s with an irreducible set of matrices must be a *scalar* matrix.

9.7 Characters

If $\mathbf{D}^{(\nu)}(R)$ and $\mathbf{D}^{(\mu)}(R)$ are related by a similarity transformation then $\mathbf{D}^{(\nu)}(R)$ gives a representation of G that is equivalent to $\mathbf{D}^{(\mu)}(R)$. These two sets of matrices are generally different, whereas their structure is the same. We wish, therefore, to answer the question: what intrinsic properties of the matrix representations are invariant under coordinate transformations?

Consider

$$\sum_i [\mathbf{CD}(R)\mathbf{C}^{-1}]_{ii} = \sum_{ikl} \mathbf{C}_{ik}\mathbf{D}_{kl}(R)\mathbf{C}_{li}^{-1}$$

$$= \sum_{kl} \delta_{kl}\mathbf{D}_{kl}(R)$$

$$= \sum_k \mathbf{D}_{kk}(R)\text{ , the trace of }\mathbf{D}(R).$$

We see that the trace, or character, is an invariant under a change of coordinate axes. We write the character as

$$\chi(R) = \sum_i \mathbf{D}_{ii}(R)$$

Equivalent representations have the same set of characters. The character of R in the representation μ is written

$$\chi^{(\mu)}(R) \text{ or } [\mu; R].$$

Now, the conjugate elements of G have the form $\mathbf{S} = \mathbf{URU}^{-1}$, and then

$$\mathbf{D}(R) = \mathbf{D}(U)\mathbf{D}(R)[\mathbf{D}(R)]^{-1}$$

therefore

$$\chi(S) = \chi(R).$$

We can describe G by giving its characters in a particular representation; all elements in a class have the same χ.

10

SOME LIE GROUPS OF TRANSFORMATIONS

We shall consider those Lie groups that can be described by a finite set of continuously varying essential parameters $a_1,...a_r$:

$$x_i' = f_i(x_1,...x_n; a_1,...a_r) = f(x; a) .$$

A set of parameters $\bar{a}$ exists that is associated with the inverse transformations:

$$x = \bar{f}(x'; \bar{a}).$$

These equations must be solvable to give the x_i's in terms of the x_i''s.

10.1 Linear groups

The *general linear group* GL(n) in n-dimensions is given by the set of equations

$$x_i' = \sum_{j=1}^{n} a_{ij}x_j \text{ , } i = 1 \text{ to } n,$$

in which $\det |a_{ij}| \neq 0$.

The group contains n^2 parameters that have values covering an infinite range. The group GL(n) is said to be *not closed.*

All linear groups with $n > 1$ are non-abelian. The group GL(n) is isomorphic to the group of $n \times n$ matrices; the law of composition is therefore matrix multiplication.

The *special linear group* of transformations SL(n) in n-dimensions is obtained from GL(n) by imposing the condition $\det| a_{ij} | = 1$. A functional relation therefore exists among the n^2 - parameters so that the number of required parameters is reduced to $(n^2 - 1)$.

10.2 Orthogonal groups

If the transformations of the general linear group GL(n) are such that

$$\sum_{i=1}^{n} x_i^{\,2} \rightarrow \text{invariant},$$

then the restricted group is called the *orthogonal group*, O(n), in n-dimensions. There are [n + n(n - 1)/2] conditions imposed on the n^2 parameters of GL(n), and therefore there are n(n - 1)/2 essential parameters of O(n).

For example, in three dimensions

$$\mathbf{x}' = \mathbf{O}\mathbf{x}\ ;\ \mathbf{O} \equiv \{\ \mathbf{O}_{3\times 3}: \mathbf{O}\mathbf{O}^{\mathrm{T}} = \mathbf{I},\ \det\mathbf{O} = 1,\ a_{ij} \quad \mathbb{R}\}$$

where

$$\mathbf{O} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

We have

$$x_1'^{2} + x_2'^{2} + x_3'^{2} = x_1^{\,2} + x_2^{\,2} + x_3^{\,2} \rightarrow \text{invariant under O(3)}.$$

This invariance imposes six conditions on the original nine parameters, and therefore O(3) is a three-parameter group.

10.3 Unitary groups

If the x_i's and the a_{ij}'s of the general linear group GL(n) are complex, and the transformations are required to leave $\mathbf{x}\mathbf{x}^{\dagger}$ invariant in the complex space, then we obtain the unitary group U(n) in n-dimensions:

$$\mathrm{U(n)} \equiv \{\ \mathbf{U}_{n\times n}: \mathbf{U}\mathbf{U}^{\dagger} = \mathbf{I},\ \det\mathbf{U} \neq 0,\ u_{ij} \quad \mathbb{C}\}.$$

There are $2n^2$ independent real parameters (the real and imaginary parts of the a_{ij}'s), and the unitary condition imposes $n + n(n-1)$ conditions on them so the group has n^2 real parameters. The unitary condition means that

$$\sum_j |a_{ij}|^2 = 1,$$

and therefore

$$|a_{ij}|^2 \leq 1 \text{ for all i, j.}$$

The parameters are limited to a finite range of values, and therefore the group U(n) is said to be *closed.*

10.4 Special unitary groups

If we impose the restriction $\det\mathbf{U} = +1$ on the unitary group U(n), we obtain the *special unitary group* SU(n) in n-dimensions:

$$SU(n) \equiv \{\mathbf{U}_{n\times n}: \mathbf{U}\mathbf{U}^\dagger = \mathbf{I}, \det\mathbf{U} = +1, u_{ij} \in \mathbf{C}\}.$$

The determinantal condition reduces the number of required real parameters to $(n^2 - 1)$. SU(2) and SU(3) are important in Modern Physics.

10.5 The group SU(2), the infinitesimal form of SU(2), and the Pauli spin matrices

The special unitary group in 2-dimensions, SU(2), is defined as

$$SU(2) \equiv \{\mathbf{U}_{2\times 2}: \mathbf{U}\mathbf{U}^\dagger = \mathbf{I}, \det\mathbf{U} = +1, u_{ij} \in \mathbf{C}\}.$$

It is a three-parameter group.

The defining conditions can be used to obtain the matrix representation in its simplest form; let

$$\mathbf{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

where a, b, c, d $\in \mathbb{C}$.

The hermitian conjugate is

$$\mathbf{U}^{\dagger} = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix},$$

and therefore

$$\mathbf{U}\mathbf{U}^{\dagger} = \begin{pmatrix} |a|^2 + |b|^2 & ac^* + bd^* \\ a^*c + b^*d & |c|^2 + |d|^2 \end{pmatrix}.$$

The unitary condition gives

$$|a|^2 + |b|^2 = |c|^2 + |d|^2 = 1,$$

and the determinantal condition gives

$$ad - bc = 1.$$

Solving these equations , we obtain

$$c = -b^*, \text{ and } d = a^*.$$

The general form of SU(2) is therefore

$$\mathbf{U} = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}.$$

We now study the infinitesimal form of SU(2); it must have the structure

$$\mathbf{U}_{\text{inf}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \delta a & \delta b \\ -\delta b^* & \delta a^* \end{pmatrix} = \begin{pmatrix} 1 + \delta a & \delta b \\ -\delta b^* & 1 + \delta a^* \end{pmatrix}.$$

The determinantal condition therefore gives

$$\det \mathbf{U}_{\text{inf}} = (1 + \delta a)(1 + \delta a^*) + \delta b \delta b^* = 1.$$

To first order in the δ's, we obtain

$$1 + \delta a^* + \delta a = 1,$$

or

$$\delta a = -\delta a^*.$$

so that

$$\mathbf{U}_{inf} = \begin{pmatrix} 1 + \delta a & \delta b \\ -\delta b^* & 1 - \delta a \end{pmatrix}.$$

The matrix elements can be written in their complex forms:

$$\delta a = i\delta\alpha/2 \, , \, \delta b = \delta\beta/2 + i\delta\gamma/2.$$

(The factor of two has been introduced for later convenience).

$$\mathbf{U}_{inf} = \begin{pmatrix} 1 + i\delta\alpha/2 & \delta\beta/2 + i\delta\gamma/2 \\ -\delta\beta/2 + i\delta\gamma/2 & 1 - i\delta\alpha/2 \end{pmatrix}.$$

Now, any 2×2 matrix can be written as a linear combination of the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

as follows

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = A \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + B \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + C \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + D \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

where

$$a = A + D, \; b = B - iC, \; c = B + iC, \text{ and } d = A - D.$$

We then have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{(a + d)}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{(b + c)}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \frac{i(b - c)}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \frac{(a - d)}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The infinitesimal form of SU(2) can therefore be written

$$\mathbf{U}_{inf} = \mathbf{I} + (i\delta\gamma/2)\sigma_1 + (i\delta\beta/2)\sigma_2 + (i\delta\alpha/2)\sigma_3 ,$$

or

$$\mathbf{U}_{inf} = \mathbf{I} + (i/2)\sum \delta\tau_j \sigma_j \,.\; j = 1 \text{ to } 3.$$

This is the Lie form.

The σ_j's are the *Pauli spin-matrices*:; they are the generators of the group SU(2):

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

They play a fundamental role in the description of spin-1/2 particles in Quantum Mechanics. (See later discussions).

10.6 Commutators of the spin matrices and structure constants

We have previously introduced the commutators of the infinitesimal generators of a Lie group in connection with their Lie Algebra. In this section, we consider the commutators of the generators of SU(2); they are found to have the symmetric forms

$$[\sigma_1, \sigma_2] = 2i\sigma_3, \; [\sigma_2, \sigma_1] = -2i\sigma_3,$$

$$[\sigma_1, \sigma_3] = -2i\sigma_2, [\sigma_3, \sigma_1] = 2i\sigma_2,$$

$$[\sigma_2, \sigma_3] = 2i\sigma_1, \; [\sigma_3, \sigma_2] = -2i\sigma_1.$$

We see that the commutator of any pair of the three matrices gives a constant multiplied by the value of the remaining matrix, thus

$$[\sigma_j, \sigma_k] = \varepsilon_{jk\ell} 2i\sigma_\ell \; .$$

where the quantity $\varepsilon_{jk\ell} = \pm 1$, depending on the permutations of the indices. ($\varepsilon_{(xy)z} = +1$, $\varepsilon_{(yx)z} = -1$ ..etc...).

The quantities $2i\varepsilon_{jk\ell}$ are the *structure constants* associated with the group.

Other properties of the spin matrices are found to be

$$\sigma_1^{\,2} = \sigma_2^{\,2} = \sigma_3^{\,2} = \mathbf{I};\ \sigma_1\sigma_2 = i\sigma_3,\ \sigma_2\sigma_3 = i\sigma_1,\ \sigma_3\sigma_1 = i\sigma_2.$$

10.7 Homomorphism of SU(2) and $O^+(3)$

We can form the matrix

$$\mathbf{P} = \mathbf{x}^T\sigma = x_j\sigma_j,\ j = 1,\ 2,\ 3$$

from the matrices

$$\mathbf{x} = [x_1, x_2, x_3] \text{ and } \sigma = [\sigma_1, \sigma_2, \sigma_3]:$$

therefore

$$\mathbf{P} = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix}.$$

We see that

$$\mathbf{P}^\dagger = (\mathbf{P}^*)^T = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix} = \mathbf{P},$$

so that **P** is hermitian.

Furthermore,

$$\mathrm{Tr}\mathbf{P} = 0,$$

and

$$\det\mathbf{P} = -(x_1^{\,2} + x_2^{\,2} + x_3^{\,2}).$$

Another matrix, **P´**, can be formed by carrying out a similarity transformation, thus

$$\mathbf{P}' = \mathbf{U}\mathbf{P}\mathbf{U}^\dagger,\ (\mathbf{U} \quad SU(2)).$$

A similarity transformation leaves both the trace and the determinant

unchanged, therefore

$$\mathrm{Tr}\mathbf{P} = \mathrm{Tr}\mathbf{P}',$$

and

$$\det\mathbf{P} = \det\mathbf{P}'.$$

However, the condition $\det\mathbf{P} = \det\mathbf{P}'$ means that

$$\mathbf{x}\mathbf{x}^{\mathrm{T}} = \mathbf{x}'\mathbf{x}'^{\mathrm{T}},$$

or

$$x_1^2 + x_2^2 + x_3^2 = x_1'^2 + x_2'^2 + x_3'^2 .$$

The transformation $\mathbf{P}' = \mathbf{U}\mathbf{P}\mathbf{U}^{\dagger}$ is therefore equivalent to a *three-dimensional* orthogonal transformation that leaves $\mathbf{x}\mathbf{x}^{\mathrm{T}}$ invariant.

10.8 Irreducible representations of SU(2)

We have seen that the basic form of the 2×2 matrix representation of the group SU(2) is

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \; a, b \in \mathbb{C}; \; |a|^2 + |b|^2 = 1.$$

Let the basis vectors of this space be

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We then have

$$\mathbf{x}_1' = \mathbf{U}\mathbf{x}_1 = \begin{pmatrix} a \\ -b^* \end{pmatrix} = a\mathbf{x}_1 - b^*\mathbf{x}_2 ,$$

and

$$\mathbf{x}_2' = \mathbf{U}\mathbf{x}_2 = \begin{pmatrix} b \\ a^* \end{pmatrix} = b\mathbf{x}_1 + a^*\mathbf{x}_2 ,$$

and therefore

$$\mathbf{x}' = \mathbf{U}^{\mathrm{t}}\mathbf{x}.$$

If we write a 2-dimensional vector in this complex space as $\mathbf{c} = [u, v]$

then the components transform under SU(2) as

$$u' = au + bv$$

and

$$v' = b^*u + a^*v ,$$

and therefore

$$\mathbf{c}' = \mathbf{U}\mathbf{c} .$$

We see that the components of the vector **c** transform *differently* from those of the basis vector **x** — the transformation matrices are the transposes of each other. The vector **c** = [u, v] in this complex space is called a spinor (Cartan, 1913).

To find an irreducible representation of SU(2) in a 3-dimensional space, we need a set of three linearly independent basis functions. Following Wigner (see bibliography), we can choose the polynomials

$$u^2, uv, \text{ and } v^2,$$

and introduce the polynomials defined by

$$f^{j=1}_{m} = \frac{u^{1+m} v^{1-m}}{\sqrt{\{(1+m)!\,(1+m)!\}}}$$

where

$$j = n/2 \text{ (the dimension of the space is } n + 1) .$$

and

$$m = j, j - 1, \ldots -j .$$

In the present case, n = 2, j = 1, and m = 0, ±1.

(The factor $1/\sqrt{\{(1+m)!\,(1-m)!\}}$ is chosen to make the representative

matrix unitary).

We have, therefore

$$f_1^{\,1} = u^2/\sqrt{2}\ ,\ f_0^{\,1} = uv, \text{ and } f_{-1}^{\,1} = v^2/\sqrt{2}.$$

A 3×3 representation of an element $\mathbf{U} \in SU(2)$ in this space can be found by defining the transformation

$$\mathbf{U}f_m^{\,1}(u, v) = f_m^{\,1}(u', v').$$

We then obtain

$$\mathbf{U}f_m^{\,1}(u, v) = \frac{(au + bv)^{1+m}(-b^*u + a^*v)^{1-m}}{\sqrt{\{(1+m)!(1-m)!\}}}\ ,\ m = 0, \pm 1,$$

so that

$$\begin{aligned}
\mathbf{U}f_1^{\,1}(u, v) &= (au + bv)^2/\sqrt{2} \\
&= (a^2u^2 + 2abuv + b^2v^2)/\sqrt{2}\ , \\
\mathbf{U}f_0^{\,1}(u, v) &= (au + bv)(-b^*u + a^*v) \\
&= -ab^*u^2 + (|a|^2 - |b|^2)uv + a^*bv^2\ ,
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{U}f_{-1}^{\,1}(u, v) &= (-b^*u + a^*v)^2/\sqrt{2} \\
&= (b^{*2}u^2 - 2a^*b^*uv + a^{*2}v^2)/\sqrt{2}\ .
\end{aligned}$$

We then have

$$\begin{pmatrix} a^2 & \sqrt{2}ab & b^2 \\ -\sqrt{2}ab^* & |a|^2 - |b|^2 & \sqrt{2}a^*b \\ b^{*2} & -\sqrt{2}a^*b^* & a^{*2} \end{pmatrix} \begin{pmatrix} f_1^{\,1} \\ f_0^{\,1} \\ f_{-1}^{\,1} \end{pmatrix} = \begin{pmatrix} f_1^{\,1'} \\ f_0^{\,1'} \\ f_{-1}^{\,1'} \end{pmatrix}$$

or

$$\boldsymbol{U}\mathbf{F} = \mathbf{F}'.$$

We find that $\boldsymbol{U}\boldsymbol{U}^{\dagger} = \mathbf{I}$ and therefore $\boldsymbol{U}$ is, indeed, unitary.

This procedure can be generalized to an (n + 1)-dimensional space as follows

Let

$$f_m^j(u, v) = \frac{u^{j+m} v^{j-m}}{\sqrt{\{(j+m)!(j-m)!\}}}, \quad m = j,\ j-1,\ \ldots -j.$$

(Note that $j = n/2 = 1/2, 1/1, 3/2, 2/1, ..$).

For a given value of j, there are $2j + 1$ linearly independent polynomials, and therefore we can form a $(2j + 1) \times (2j + 1)$ representative matrix of an element **U** of SU(2):

$$\mathbf{U} f_m^j(u, v) = f_m^j(u', v').$$

The details of this general case are given in Wigner's classic text. He demonstrates the irreducibility of the $(2j + 1)$-dimensional representation by showing that any matrix **M** which commutes with $\mathbf{U}^j$ for all a, b such that $|a|^2 + |b|^2 = 1$ must necessarily be a *constant* matrix, and therefore, by Schur's lemma, $\mathbf{U}^j$ is an irreducible representation.

10.9 Representations of rotations and the concept of tensors

We have discussed 2- and 3-dimensional representations of the orthogonal group O(3) and their connection to angular momentum operators. Higher-dimensional representations of the orthogonal group can be obtained by considering a 2-index quantity , T_{ij} — a tensor — that consists of a set of 9 elements that transform under a rotation of the coordinates as follows:

$$T_{ij} \rightarrow T_{ij}' = R_{i\ell} R_{jm} T_{\ell m} \text{ (sum over repeated indices 1, 2, 3).}$$

If $T_{ij} = T_{ji}$ (T_{ij} is symmetric), then this symmetry is an *invariant* under rotations; we have

$$T_{ji}' = R_{j\ell} R_{im} T_{\ell m} = R_{jm} R_{i\ell} T_{m\ell} = R_{i\ell} R_{jm} T_{\ell m} = T_{ij}' .$$

If $TrT_{ij} = 0$, then so is TrT_{ij}', for

$$T_{ii}' = R_{i\ell} R_{im} T_{\ell m} = (R^T R)_{\ell m} T_{\ell m} = \delta_{\ell m} T_{\ell m} = T_{\ell\ell} = 0.$$

The components of a symmetric traceless 2-index tensor contains 5 members so that the transformation $T_{ij} \rightarrow T_{ij}' = R_{i\ell} R_{jm} T_{\ell m}$ defines a new representation of them of dimension 5.

Any tensor $\bar{T}_{ij}$ can be written

$$\bar{T}_{ij} = (\bar{T}_{ij} + \bar{T}_{ji})/2 + (\bar{T}_{ij} - \bar{T}_{ji})/2 ,$$

and we have

$$\bar{T}_{ij} = (\bar{T}_{ij} + \bar{T}_{ji})/2 = (\bar{T}_{ij} - (\delta_{ij} T_{\ell\ell})/3) + (\delta_{ij} T_{\ell\ell})/3 .$$

The decomposition of the tensor $\bar{T}_{ij}$ gives any 2-index tensor in terms of a sum of a single component, proportional to the identity, a set of 3 independent quantities combined in an anti-symmetric tensor $(\bar{T}_{ij} - \bar{T}_{ji})/2$, and a set of 5 independent components of a symmetric traceless tensor. We write the *dimensional* equation

$$9 = 1 \quad 3 \quad 5 .$$

This is as far as it is possible to go in the process of decomposition: no other subsets of 2-index tensors can be found that preserve their identities under the defining transformation of the coordinates. *Representations with no subsets of tensors that preserve their identities under the defining rotations of tensors are irreducible representations.*

We shall see that the decomposition of tensor products into symmetric and anti-symmetric parts is important in the Quark Model of elementary particles.

The representations of the orthogonal group O(3) are found to be important in defining the *intrinsic spin* of a particle. The dynamics of a particle of finite mass can always be descibed in its rest frame (all inertial frames are equivalent!), and therefore the particle can be characterized by rotations. All known particles have dynamical states that can be described in terms of the tensors of some irreducible representation of O(3). If the dimension of the irrep is (2j + 1) then the particle spin is found to be proportional to j. In Particle Physics, irreps with values of j = 0, 1, 2,... and with j = 1/2, 3/2, ... are found that correspond to the fundamental bosons and fermions, respectively.

The three dimensional orthogonal group SO(3) (det = +1) and the two dimensional group SU(2) have the same Lie algebra. In the case of the group SU(2), the (2j + 1)-dimensional representations are allowed for both integer and half -integer values of j, whereas, the representations of the group SO(3) are limited to integer values of j. Since all the representations are allowed in SU(2), it is called the *covering* group. We note that rotations through φ and $\varphi + 2\pi$ have different effects on the 1/2-integer representations, and therefore they are (spinor) transfomations associated with SU(2).

11

THE GROUP STRUCTURE OF LORENTZ TRANSFORMATIONS

The square of the invariant interval s, between the origin [0, 0, 0, 0] of a spacetime coordinate system and an arbitrary event $x^\mu = [x^0, x^1, x^2, x^3]$ is, in index notation

$$s^2 = x^\mu x_\mu = x'^\mu x'_\mu \text{ , (sum over } \mu = 0, 1, 2, 3).$$

The lower indices can be raised using the metric tensor

$$\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1),$$

so that

$$s^2 = \eta_{\mu\nu} x^\mu x^\nu = \eta_{\mu\nu} x'^\mu x'^\nu \text{ , (sum over } \mu \text{ and } \nu).$$

The vectors now have contravariant forms.

In matrix notation, the invariant is

$$s^2 = \mathbf{x}^T \eta \mathbf{x} = \mathbf{x}'^T \eta \mathbf{x}' .$$

(The transpose must be written explicitly).

The primed and unprimed column matrices (contravariant vectors) are related by the Lorentz matrix operator, **L**

$$\mathbf{x}' = \mathbf{L}\mathbf{x} .$$

We therefore have

$$\mathbf{x}^T \eta \mathbf{x} = (\mathbf{L}\mathbf{x})^T \eta (\mathbf{L}\mathbf{x})$$

$$= \mathbf{x}^T \mathbf{L}^T \eta \mathbf{L}\mathbf{x} .$$

The **x**'s are arbitrary, therefore

$$\mathbf{L}^T \eta \mathbf{L} = \eta.$$

This is the *defining* property of the Lorentz transformations.

The set of all Lorentz transformations is the set $\boldsymbol{L}$ of all 4 × 4 matrices that satisfies the defining property

$$\boldsymbol{L} = \{\mathbf{L}: \mathbf{L}^{T}\eta\mathbf{L} = \eta;\ \mathbf{L}\text{: all } 4 \times 4 \text{ real matrices};$$

$$\eta = \text{diag}(1, -1, -1, -1\}.$$

(Note that each $\boldsymbol{L}$ has 16 (independent) real matrix elements, and therefore belongs to the 16-dimensional space, R^{16}).

11.1 The group structure of $\boldsymbol{L}$

Consider the result of two successive Lorentz transformations $\mathbf{L}_1$ and $\mathbf{L}_2$ that transform a 4-vector $\mathbf{x}$ as follows

$$\mathbf{x} \rightarrow \mathbf{x}' \rightarrow \mathbf{x}''$$

where

$$\mathbf{x}' = \mathbf{L}_1\mathbf{x}\ ,$$

and

$$\mathbf{x}'' = \mathbf{L}_2\mathbf{x}'.$$

The resultant vector $\mathbf{x}''$ is given by

$$\begin{aligned}\mathbf{x}'' &= \mathbf{L}_2(\mathbf{L}_1\mathbf{x}) \\ &= \mathbf{L}_2\mathbf{L}_1\mathbf{x} \\ &= \mathbf{L}_c\mathbf{x}\end{aligned}$$

where

$$\mathbf{L}_c = \mathbf{L}_2\mathbf{L}_1\ (\mathbf{L}_1 \text{ followed by } \mathbf{L}_2).$$

If the combined operation $\mathbf{L}_c$ is always a Lorentz transformation then it must satisfy

$$\mathbf{L}_c^{\,T}\eta\mathbf{L}_c = \eta\ .$$

We must therefore have

$$(\mathbf{L}_2\mathbf{L}_1)^T\eta(\mathbf{L}_2\mathbf{L}_1) = \eta$$

or

$$\mathbf{L}_1^{\,T}(\mathbf{L}_2^{\,T}\eta\mathbf{L}_2)\mathbf{L}_1 = \eta$$

so that

$$\mathbf{L}_1^{\,T}\eta\mathbf{L}_1 = \eta, \quad (\mathbf{L}_1, \mathbf{L}_2 \quad \mathbf{L})$$

therefore

$$\mathbf{L}_c = \mathbf{L}_2\mathbf{L}_1 \quad \mathbf{L}\ .$$

Any number of successive Lorentz transformations may be carried out to give a resultant that is itself a Lorentz transformation.

If we take the determinant of the defining equation of **L**,

$$\det(\mathbf{L}^T\eta\mathbf{L}) = \det\eta$$

we obtain

$$(\det\mathbf{L})^2 = 1 \quad (\det\mathbf{L} = \det\mathbf{L}^T)$$

so that

$$\det\mathbf{L} = \pm 1.$$

Since the determinant of **L** is not zero, an inverse transformation $\mathbf{L}^{-1}$ exists, and the equation $\mathbf{L}^{-1}\mathbf{L} = \mathbf{I}$, the identity, is always valid.

Consider the inverse of the defining equation

$$(\mathbf{L}^T\eta\mathbf{L})^{-1} = \eta^{-1}\ ,$$

or

$$\mathbf{L}^{-1}\eta^{-1}(\mathbf{L}^{\mathrm{T}})^{-1} = \eta^{-1} .$$

Using $\eta = \eta^{-1}$, and rearranging, gives

$$\mathbf{L}^{-1}\eta(\mathbf{L}^{-1})^{\mathrm{T}} = \eta .$$

This result shows that the inverse $\mathbf{L}^{-1}$ is always a member of the set ***L***.

We therefore see that

1. If $\mathbf{L}_1$ and $\mathbf{L}_2$ $\quad \boldsymbol{L}$, then $\mathbf{L}_2\,\mathbf{L}_1$ $\quad \boldsymbol{L}$

2. If $\mathbf{L}$ $\quad \boldsymbol{L}$, then $\mathbf{L}^{-1}$ $\quad \boldsymbol{L}$

3. The identity $\mathbf{I} = \mathrm{diag}(1, 1, 1, 1)$ $\quad \boldsymbol{L}$

and

4. The matrix operators **L** obey associativity.

The set of all Lorentz transformations therefore forms a *group*.

11.2 The rotation group, revisited

Spatial rotations in two and three dimensions are Lorentz transformations in which the time-component remains unchanged.

Let R be a real 3×3 matrix that is part of a Lorentz transformation with a constant time-component. In this case, the defining property of the Lorentz transformations leads to

$$R^{\mathrm{T}}R = \mathbf{I} \text{, the identity matrix, } \mathrm{diag}(1,1,1).$$

This is the defining property of a three-dimensional orthogonal matrix

If $\mathbf{x} = [x_1, x_2, x_3]$ is a three-vector that is transformed under R to give $\mathbf{x}'$ then

$$\mathbf{x}'^{T}\mathbf{x}' = \mathbf{x}^{T} R^{T} R\,\mathbf{x}$$

$$= \mathbf{x}^{T}\mathbf{x} = x_1^{\,2} + x_2^{\,2} + x_3^{\,2}$$

$$= \text{invariant under } R\,.$$

The action of R on any three-vector preserves length. The set of all 3×3 orthogonal matrices is denoted by $\mathbf{O}(3)$,

$$\mathbf{O}(3) = \{\, R :\; R^{T} R = \mathbf{I},\; r_{ij} \quad \mathbf{R} \,\}.$$

The elements of this set satisfy the four group axioms.

The group $\mathbf{O}(3)$ can be split into two parts that are said to be *disconnected*:: one with $\det R = +1$ and the other with $\det R = -1$. The two parts are written

$$\mathbf{O}^{+}(3) \;=\; \{\, R : \det R = +1 \,\}$$

and

$$\mathbf{O}^{-}(3) \;=\; \{\, R : \det R = -1 \,\}\;.$$

If we define the *parity operator* , **P**, to be the operator that reflects all points in a 3-dimensional cartesian system through the origin then

$$\mathbf{P} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The two parts of $\mathbf{O}(3)$ are related by the operator **P**:

if $R \quad \mathbf{O}^{+}(3)$ then $\mathbf{P}R \quad \mathbf{O}^{-}(3)$,

and

if $R' \in \mathbf{O}^-(3)$ then $\mathbf{P}R' \in \mathbf{O}^+(3)$.

We can therefore consider only that part of $\mathbf{O}(3)$ that is a group, namely $\mathbf{O}^+(3)$, together with the operator $\mathbf{P}$.

11.3 Connected and disconnected parts of the Lorentz group

We have shown, previously, that every Lorentz transformation, $\mathbf{L}$, has a determinant equal to ±1. The matrix elements of $\mathbf{L}$ change continuously as the relative velocity changes continuously. It is not possible, however, to move continuously in such a way that we can go from the set of transformations with $\det\mathbf{L} = +1$ to those with $\det\mathbf{L} = -1$; we say that the set $\{\mathbf{L}: \det\mathbf{L} = +1\}$ is *disconnected* from the set $\{\mathbf{L}: \det\mathbf{L} = -1\}$.

If we write the Lorentz transformation in its component form

$$\mathbf{L} \rightarrow L^{\mu}{}_{\nu}$$

where $\mu = 0,1,2,3$ labels the rows, and $\nu = 0,1,2,3$ labels the columns then the time component $L^0{}_0$ has the values

$$L^0{}_0 \geq +1 \text{ or } L^0{}_0 \leq -1.$$

The set of transformations can therefore be split into four disconnected parts, labelled as follows:

$$\{\mathbf{L}^{\uparrow}{}_{+}\} = \{\mathbf{L}: \det\mathbf{L} = +1,\ L^0{}_0 \geq +1\}$$

$$\{\mathbf{L}^{\uparrow}{}_{-}\} = \{\mathbf{L}: \det L = -1,\ L^0{}_0 \geq +1\}$$

$$\{L^{\downarrow}{}_{+}\} = \{\mathbf{L}: \det\mathbf{L} = +1,\ L^0{}_0 \leq -1\},$$

and

$$\{\mathbf{L}^{\downarrow}_{-}\} = \{\mathbf{L}: \det\mathbf{L} = -1,\ L^{0}_{0} \leq -1\}.$$

The identity is in $\{\mathbf{L}^{\uparrow}_{+}\}$.

11.4 Parity, time-reversal and orthochronous transformations

Two discrete Lorentz transformations are

i) the parity transformation

$$\mathbf{P} = \{\mathbf{P}: \mathbf{r} \rightarrow -\mathbf{r},\ t \rightarrow t\}$$

$$= \mathrm{diag}(1, -1, -1, -1),$$

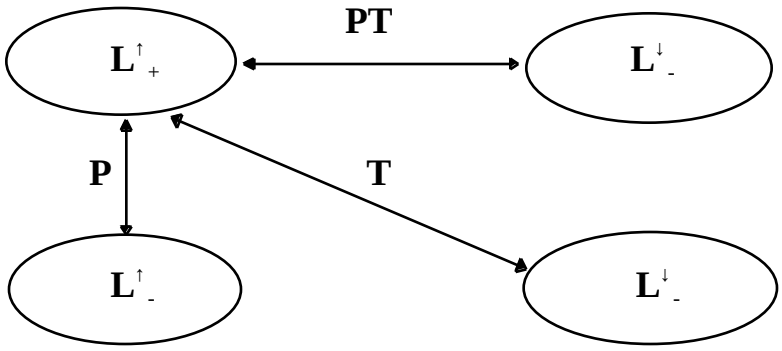
and

ii) the time-reversal transfprmation

$$\mathbf{T} = \{\mathbf{T}: \mathbf{r} \rightarrow \mathbf{r},\ t \rightarrow -t\}$$

$$= \mathrm{diag}(-1, 1, 1, 1\}.$$

The disconnected parts of $\{\mathbf{L}\}$ are related by the transformations that involve **P**, **T**, and **PT**, as shown:



Connections between the disconnected parts of Lorentz transformations

The proper orthochronous transformations are in the group $\mathbf{L}^{\uparrow}_{+}$. We see that it is not necessary to consider the complete set $\{\mathbf{L}\}$ of Lorentz

transformations — we need consider only that subset $\{\mathbf{L}^{\uparrow}_{+}\}$ that forms a group by itself, and either **P**, **T**, or **PT** combined. Experiments have shown clear violations under the parity transformation, **P** and violations under **T** have been inferred from experiment and theory, combined. However, not a single experiment has been carried out that shows a violation of the proper orthochronous transformations, $\{\mathbf{L}^{\uparrow}_{+}\}$.

12

ISOSPIN

Particles can be distinguished from one another by their intrinsic properties: mass, charge, spin, parity, and their electric and magnetic moments. In our on-going quest for an understanding of the true nature of the fundamental particles, and their interactions, other intrinsic properties, with names such as "isospin" and "strangeness", have been discovered. The intrinsic properties are defined by quantum numbers; for example, the quantum number *a* is defined by the eigenvalue equation

$$\mathbf{A}\varphi = a\,\varphi$$

where **A** is a linear operator, φ is the wavefunction of the system in the zero-momentum frame, and *a* is an eigenvalue of **A**.

In this chapter, we shall discuss the first of these new properties to be introduced, namely, *isospin.*

The building blocks of nuclei are protons (positively charged) and neutrons (neutral). Numerous experiments on the scattering of protons by protons, and protons by neutrons, have shown that the nuclear forces

between pairs have the same strength, provided the angular momentum and spin states are the same. These observations form the basis of an important concept — the *charge-independence of the nucleon-nucleon force.* (Corrections for the coulomb effects in proton-proton scattering must be made). The origin of this concept is found in a new *symmetry principle.* In 1932, Chadwick not only identified the neutron in studying the interaction of alpha-particles on beryllium nuclei but also showed that its mass is almost equal to the mass of the proton. (Recent measurements give

$$\text{mass of proton} = 938\ 27231(28)\ \text{MeV}/c^2$$

and

$$\text{mass of neutron} = 939\ 56563(28)\ \text{MeV}/c^2)$$

Within a few months of Chadwick's discovery, Heisenberg introduced a theory of nuclear forces in which he considered the neutron and the proton to be two "states" of the same object — the nucleon. He introduced an intrinsic variable, later called isospin, that permits the charge states (+, 0) of the nucleons to be distinguished. This new variable is needed (in addition to the traditional space-spin variables) in the description of nucleon-nucleon scattering.

In nuclei, protons and neutrons behave in a remarkably symmetrical way: the binding energy of a nucleus is closely proportional to the number of neutrons and protons, and in light nuclei (mass number <40), the number of neutrons can be equal to the number of protons.

Before discussing the isospin of particles and nuclei, it is necessary to introduce an *extended* Pauli Exclusion Principle. In its original form, the Pauli Exclusion Principle was introduced to account for features in the observed spectra of atoms that could not be understood using the then current models of atomic structure:

no two electrons in an atom can exist in the same quantum state defined by the quantum numbers n, ℓ, m_ℓ, m_s where n is the principal quantum number, ℓ is the orbital angular momentum quantum number, m_ℓ is the magnetic quantum number, and m_s is the spin quantum number.

For a system of N particles, the complete wavefunction is written as a product of single-particle wavefunctions

$$\Psi(1, 2, \ldots N) = \psi(1)\psi(2)\ldots\psi(N).$$

Consider this form in the simplest case — for two identical particles. Let one be in a state labelled Ψ_a and the other in a state Ψ_b. For identical particles, it makes no difference to the probability density $|\Psi|^2$ of the 2-particle system if the particles are exchanged:

$$|\Psi(1, 2)|^2 = |\Psi(2, 1)|^2 \text{, (the } \Psi\text{'s are not measurable)}$$

so that, either

$$\Psi(2, 1) = \Psi(1, 2) \text{ (symmetric)}$$

or

$$\Psi(2, 1) = -\Psi(1, 2) \text{ (anti-symmetric).}$$

Let

$$\Psi_I = \psi_a(1)\psi_b(2) \text{ (1 an a, 2 in b)}$$

and

$$\Psi_{II} = \psi_a(2)\psi(1) \text{ (2 in a, 1 in b).}$$

The two particles are indistinguishable, therefore we have no way of knowing whether Ψ_I or Ψ_{II} describes the system; we postulate that the system spends 50% of its time in Ψ_I and 50% of its time in Ψ_{II}. The two-particle system is considered to be a *linear combination* of Ψ_I and Ψ_{II}:

We have, therefore, either

$$\Psi_{symm} = (1/\sqrt{2})\{\psi_a(1)\psi_b(2) + \psi_a(2)\psi_b(1)\} \text{ (BOSONS)}$$

or

$$\Psi_{antisymm} = (1/\sqrt{2})\{\psi_a(1)\psi_b(2) - \psi_a(2)\psi_b(1)\} \text{ (FERMIONS)}.$$

(The coefficient $(1/\sqrt{2})$ normalizes the sum of the squares to be 1).

Exchanging $1 \leftrightarrow 2$ leaves Ψ_{symm} unchanged, whereas exchanging particles $1 \leftrightarrow 2$ reverses the sign of $\Psi_{antisymm}$.

If two particles are in Ψ_S, both particles can exist in the same state with a = b. If two particles are in Ψ_{AS}, and a = b, we have $\Psi_{AS} = 0$ — they cannot exist in the same quantum state. Electrons (fermions, spin = $(1/2)\hbar$) are described by anti-symmetric wavefunctions.

We can now introduce a more general Pauli Exclusion Principle. Write the nucleon wavefunction as a product:

$$\Psi(\chi, q) = \psi(\chi)\varphi_N(q),$$

where

$$\chi = \chi(\mathbf{r}, s)$$

in which **r** is the space vector, s is the spin, and q is a charge or isospin label.

For two nucleons, we write

$$\Psi(\chi_1, q_1; \chi_2, q_2),$$

for two protons:

$$\Psi_{2p} = \psi_1(\chi_1, \chi_2)\varphi_N(p_1)\varphi_N(p_2),$$

for two neutrons:

$$\Psi_{2n} = \psi_2(\chi_1, \chi_2)\varphi_N(n_1)\varphi_N(n_2),$$

and for an n-p pair:

$$\Psi_{np} = \psi_3(\chi_1, \chi_2)\varphi_N(p_1)\varphi_N(n_2)$$

or

$$= \psi_4(\chi_1, \chi_2)\varphi_N(n_1)\varphi_N(p_2).$$

If we regard the proton and neutron as different states of the same object, labelled by the "charge or isospin coordinate", q, we must extend the Pauli principle to cover the new coordinate: the total wavefunction is then

$$\Psi(\chi_1, q_1; \chi_2, q_2) = -\Psi(\chi_2, q_2; \chi_1, q_1)\ .$$

It must be anti-symmetric under the full exchange.

For a 2p- or a 2n-pair, the exchange $q_1 \leftrightarrow q_2$ is symmetrical, and therefore the space-spin part must be anti-symmetrical.

For an n-p pair, the symmetric (S) and anti-symmetric (AS) "isospin" wavefunctions are

I) $\Phi_S = (1/\sqrt{2})\{\varphi_N(p_1)\varphi_N(n_2) + \varphi_N(n_1)\varphi_N(p_2)\}$

(symmetric under $q_1 \leftrightarrow q_2$),

and therefore the space-spin part is anti-symmetrical,

II) $\Phi_{AS} = (1/\sqrt{2})\{\varphi_N(p_1)\varphi_N(n_2) - \varphi_N(n_1)\varphi_N(p_2)\}$

(anti-symmetric under $q_1 \leftrightarrow q_2$),

and therefore the space-spin part is symmetrical.

We shall need these results in later discussions of the symmetric and anti-symmetric properties of quark systems.

12.1 Nuclear β-decay

Nuclei are bound states of neutrons and protons. If the number of protons in a nucleus is Z and the number of neutrons is N then the mass number of the nucleus is A = N + Z. Some nuclei are naturally unstable. A possible mode of decay is by the emission of an electron (this is β-decay — a process that typifies the fundamental "weak interaction").

We write the decay as

$$^{A}_{Z}X_N \rightarrow {}^{A}_{Z+1}X_{N-1} + e^{-1} + \bar{\nu}_e \quad (\beta^-\text{-decay})$$

or, we can have

$$^{A}_{Z}X_N \rightarrow {}^{A}_{Z-1}X_{N-1} + e^{+} + \nu_e \quad (\beta^+ \text{ - decay}).$$

A related process is that of electron capture of an orbital electron that is sufficiently close to the positively charged nucleus:

$$e^- + {}^{A}_{Z}X_N \rightarrow {}^{A}_{Z+1}X_{N+1} + \nu_e.$$

Other related processes are

$$\bar{\nu}_e + {}^{A}_{Z}X_N \rightarrow {}^{A}_{Z-1}X_{N-1} + e^+$$

and

$$\nu_e + {}^{A}_{Z}X_N \rightarrow {}^{A}_{Z+1}X_{N-1} + e^- .$$

The decay of the *free* proton has not been observed at the present time. The experimental limit on the half-life of the proton is $> 10^{31}$ years! Many current theories of the microstructure of matter predict that the proton decays. If, however, the life-time is $> 10^{32}$ - 10^{33} years then there is no realistic possibility of observing the decay directly (The limit is set by Avogadro's number and the finite number of protons that can be assembled in a suitable experimental apparatus).

The fundamental β-decay is that of the free neutron, first observed in 1946. The process is

$$n^0 \rightarrow p^+ + e^- + \bar{\nu}_e{}^0 , \; t_{1/2} = 10\ 37 \pm 0\ 19 \text{ minutes.}$$

This measured life-time is of fundamental importance in Particle Physics and in Cosmology.

Let us set up an algebraic description of the β-decay process, recognizing that we have a 2-state system in which the transformation $p \leftrightarrow n$ occurs:

In the β^--decay of a free neutron

$$n \rightarrow p^+ + e^- + \bar{\nu}_e,$$

and in the β^+-decay of a proton, *bound in a nucleus,*

$$p \rightarrow n + e^+ + \nu_e .$$

12.2 Isospin of the nucleon

The spontaneous transformations $p \leftrightarrow n$ observed in β-decay lead us to introduce the operators $\tau_{\pm}$ that transform $p \leftrightarrow n$:

$$\tau_+ \varphi_n = \varphi_p\,, \quad \tau_+ \varphi_p = 0, \text{ (eliminates a proton)}$$

and

$$\tau_- \varphi_p = \varphi_n\,, \quad \tau_- \varphi_n = 0, \text{ (eliminates a neutron).}$$

Since we are dealing with a two-state system, we choose the "isospin" parts of the proton and neutron wavefunctions to be

$$\varphi(p) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \varphi(n) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

in which case the operators must have the forms:

$$\tau_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } \tau_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

They are singular and non-hermitian.

We have, for example

$$\tau_+ \varphi_n = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \varphi_n \rightarrow \varphi_p,$$

and

$$\tau_+ \varphi_p = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (\tau_+ \text{ removes a proton}).$$

To make the present algebraic description analogous to the two-state system of the intrinsic spin of the electron, we introduce linear

combinations of the $\tau_{\pm}$:

$$\tau_1 = \tau_+ + \tau_- = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_1, \text{ a Pauli matrix,}$$

and

$$\tau_2 = i(\tau_- - \tau_+) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \sigma_2, \text{ a Pauli matrix.}$$

A third operator that is diagonal is, as expected

$$\tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \sigma_3, \text{ a Pauli matrix.}$$

The three operators $\{\tau_1, \tau_2, \tau_3\}$ therefore obey the commutation relations

$$[\tau_j/2, \tau_k/2] = i\varepsilon_{jk\ell}\tau_\ell/2 ,$$

where the factor of(1/2) is introduced because of the 2:1 homomorphism between SU(2) and $O^+(3)$: the vector operator

$$\mathbf{t} = \tau/2$$

is called the *isospin operator of the nucleon*.

To classify the isospin states of the nucleon we may use the projection of **t** on the 3rd axis, t_3. The eigenvalues, t_3, of $\mathbf{t}_3$ correspond to the proton ($t_3 = +1/2$) and neutron ($t_3 = -1/2$) states. The nucleon is said to be an *isospin doublet* with isospin quantum number t = 1/2. (The number of states in the multiplet is $2t + 1 = 2$ for t = 1/2).

The charge, Q_N of the nucleon can be written in terms of the isospin quantum numbers:

$$Q_N = q(t_3 + (1/2)) = q \text{ or } 0,$$

where q is the proton charge. (It is one of the great unsolved problems of Particle Physics to understand why the charge on the proton is equal to the charge on the electron).

12.3 Isospin in nuclei.

The concept of isospin, and of rotations in isospin space, associated with individual nucleons can be applied to nuclei — systems of many nucleons in a bound state.

Let the isospin of the ith-nucleon be $\mathbf{t}_i$, and let $\mathbf{t}_i = \tau_i /2$. The operator of a system of A nucleons is defined as

$$\mathbf{T} = \sum^{A}_{i=1} \mathbf{t}_i = \sum^{A}_{i=1} \tau_i/2 .$$

The eigenvalue of T_3 of the isospin operator $\mathbf{T}_3$ is the sum of the individual components

$$T_3 = \sum^{A}_{i=1} t_{3i} = \sum^{A}_{i=1} \tau_{3i}/2$$
$$= (Z - N)/2 .$$

The charge, Q_N of a nucleus can be written

$$Q_N = q\sum^{A}_{i=1} (\tau_{3i} + 1)/2$$
$$= q(T_3 + A/2) .$$

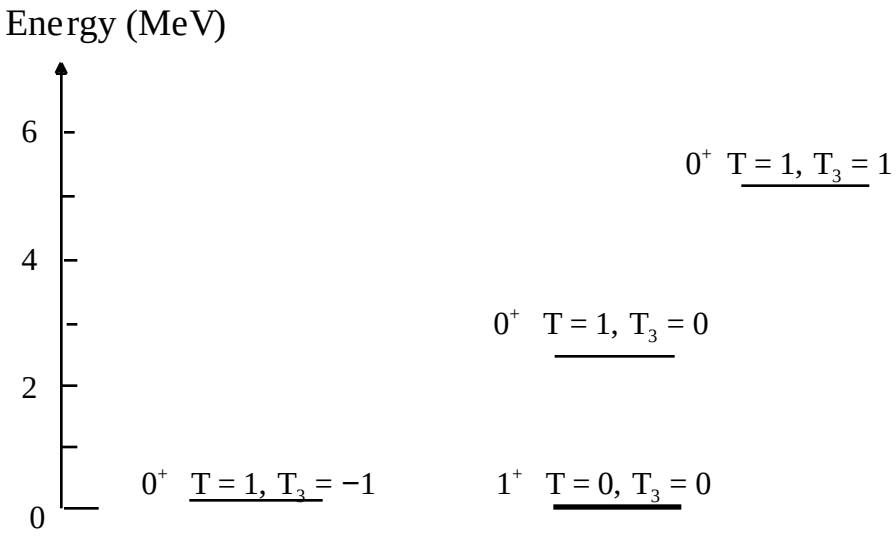
For a given eigenvalue T of the operator **T**, the state is (2T + 1)-fold degenerate. The eigenvalues T_3 of $\mathbf{T}_3$ are

$$T_3 = -T, -T + 1, ...0, ...T + 1, T .$$

If the Hamiltonian **H** of the nucleus is charge-independent then

$$[\mathbf{H}, \mathbf{T}] = 0.$$

and **T** is said to be a good quantum number. In light nuclei, where the isospin-violating coulomb interaction between pairs of protons is a small effect, the concept of isospin is particularly useful. The study of isospin effects in nuclei was first applied to the observed properties of the lowest-lying states in the three nuclei with mass number A = 14: ^{14}C, ^{14}N, and ^{14}O. The relative energies of the states are shown in the following diagram:



An isospin singlet (T = 0) and an isospin triplet (T = 1) in the A = 14 system. In the absence of the coulomb interaction, the three T = 1 states would be degenerate.

The spin and parity of the ground state of ^{14}C, the first excited state of ^{14}N and the ground state of ^{14}O are measured to be 0^+; these three states are characterized by T = 1. The ground state of ^{14}N has spin and parity 1^+; it is an isospin singlet (T = 0).

12.4 Isospin and mesons

We have seen that it is possible to classify the charge states of nucleons and nuclear isobars using the concept of isospin, and the algebra of SU(2). It will be useful to classify other particles, including field particles (quanta) in terms of their isospin.

Yukawa (1935), first proposed that the strong nuclear force between a pair of nucleons is carried by *massive* field particles called *mesons.* Yukawa's method was a masterful development of the theory of the electromagnetic field to include the case of a massive field particle. If ψ_π is the "meson wavefunction" then the Yukawa differential equation for the meson is

$$\partial_\mu\partial^\mu\ \psi_\pi + (E^0/\hbar c)^2\psi_\pi\ =\ 0.$$

where

$$\partial_\mu\partial^\mu\ =\ (1/c^2)\partial^2/\partial t^2 - \quad ^2 .$$

The r-dependent (spatial) form of ^{2} is

$$^2\ \rightarrow\ (1/r^2)d/dr(r^2d/dr)$$

The static (time-independent) solution of this equation is readily checked to be

$$\Psi(r)\ =\ (-g^2/r)\exp(-r/r_N)$$

where

$$r_N = \hbar/m_\pi c = \hbar c/m_\pi c^2 = \hbar c/E_\pi^{\ 0},$$

so that

$$1/r_N^2 = (E_\pi^0/\hbar c)^2$$

The "range of the nuclear force" is defined by the condition

$$r = r_N = \hbar/m_\pi c \approx 2 \times 10^{-13} \text{ cm.}$$

This gives the mass of the meson to be close to the measured value. It is important to note that the "range of the force" 1/(mass of the field quantum). In the case of the electromagnetic field, the mass of the field quantum (the photon) is zero, and therefore the force has an infinite range.

The mesons come in three charge states: +, −, and 0. The mesons have intrinsic spins equal to zero (they are field particles and therefore they are bosons), and their rest energies are measured to be

$$E_{\pi^\pm}^0 = 139\ 5 \text{ MeV, and } E_{\pi^0}^0 = 135\ 6 \text{ MeV.}$$

They are therefore considered to be members of an isospin triplet:

$$t = 1,\ t_3 = \pm 1,\ 0.$$

In Particle Physics, it is the custom to designate the isospin quantum number by I, we shall follow this convention from now on.

The third component of the isospin is an additive quantum number. The combined values of the isospin projections of the two particles, one with isospin projection $I_3^{(1)}$, and the other with $I_3^{(2)}$, is

$$I_3^{(1+2)} = I_3^{(1)} + I_3^{(2)}.$$

Their isospins combine to give states with different numbers in each multiplet. For example, in pion (meson)-nucleon scattering

$$\pi + N \rightarrow \text{states with } I_3^{(1+2)} = (3/2) \text{ or } (1/2).$$

These values are obtained by noting that

$$I_\pi^{(1)} = 1, \text{ and } I_N^{(2)} = 1/2, \text{ so that}$$

$$I_{3\pi}^{(1)} + I_{3N}^{(2)} = (\pm 1, 0) + (\pm 1/2)$$

$$= (3/2), \text{ an isospin quartet, or } (1/2), \text{ an isospin doublet.}$$

Symbolically , we write

$$3 \quad 2 = 4 \quad 2.$$

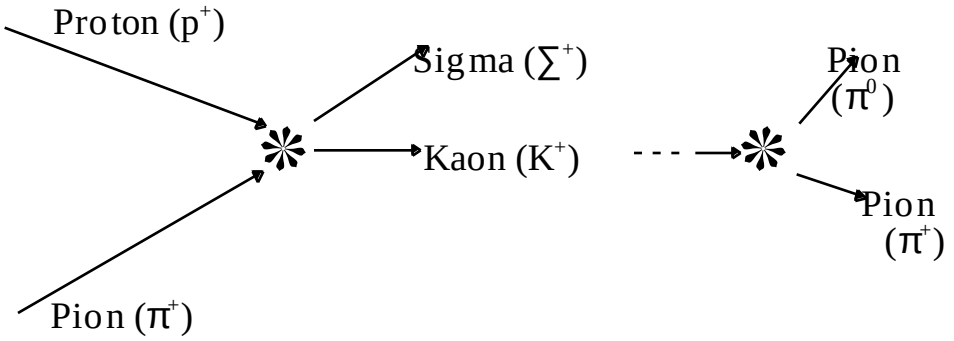
(This is the rule for forming the product $(2I_3^{(1)} + 1) \quad (2I_3^{(2)} + 1)$.

13

GROUPS AND THE STRUCTURE OF MATTER

13.1 Strangeness

In the early 1950's, our understanding of the ultimate structure of matter seemed to be complete. We required neutrons, protons, electrons and neutrinos, and mesons and photons. Our optimism was short-lived. By 1953, excited states of the nucleons, and more massive mesons, had been discovered. Some of the new particles had completely unexpected properties; for example, in the interaction between protons and π-mesons (pions) the following decay mode was observed:



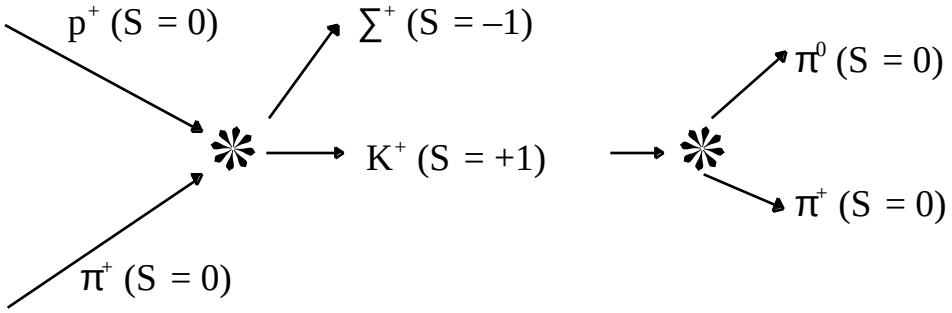
Initial interaction lasts ~10^{-23} seconds

(*Strong* force acting)

Final decay takes ~10^{-10} seconds

(*Weak* force acting)

Gell-Mann, and independently Nishijima, proposed that the kaons (heavy mesons) were endowed with a new intrinsic property not affected by the strong force. Gell-Mann called this property "strangeness". Strangeness is conserved in the strong interactions but changes in the weak interactions. The Gell-Mann - Nishijima interpretation of the strangeness-changing involved in the proton-pion interaction is



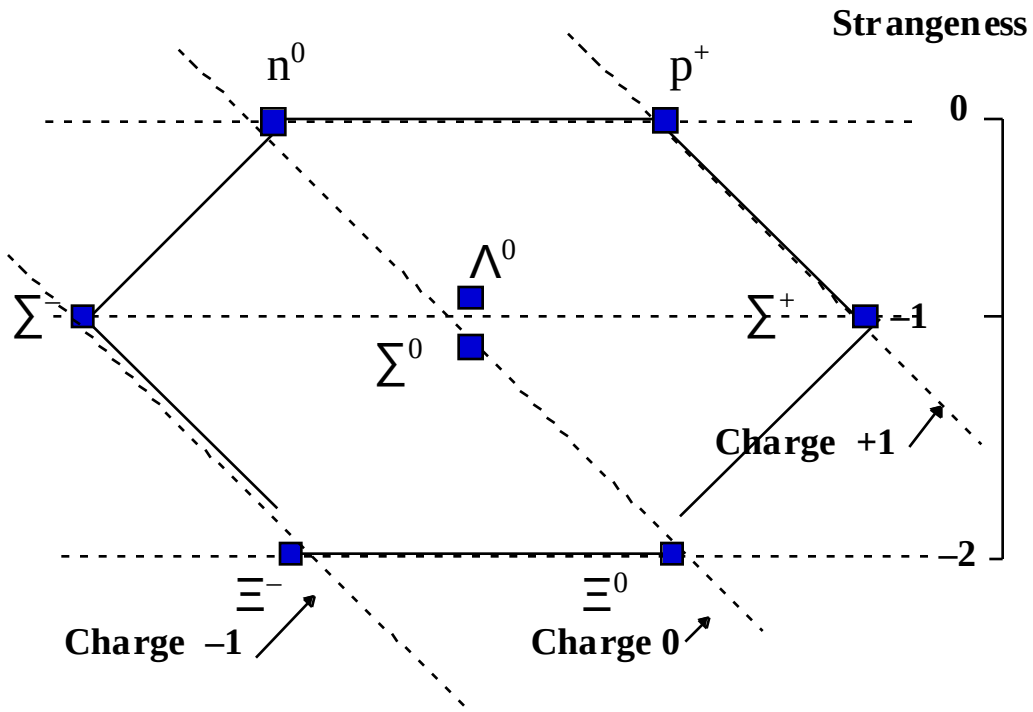
$\Delta S = 0$ $\Delta S = 1$

In the strong part of the interaction, there is no change in the number defining the strangeness, whereas in the weak part, the strangeness changes by one unit. Having defined the values of S for the particles in this

interaction, they are defined forever. All subsequent experiments involving these objects have been consistent with the original assignments.

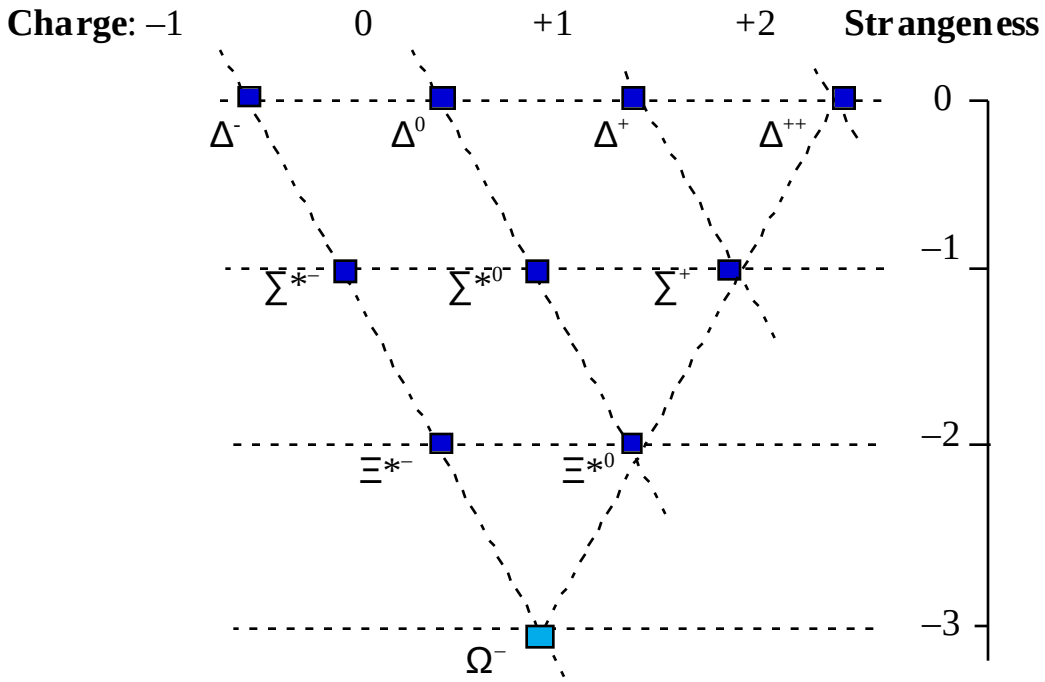
13.2 Particle patterns

In 1961, Gell-Mann, and independently Ne'eman, introduced a scheme that classified the strongly interacting particles into family groups. They were concerned with the inclusion of "strangeness" in their theory, and therefore they studied the arrangements of particles in an *abstract space* defined by their electric charge and strangeness. The common feature of each family was chosen to be their intrinsic spin; the family of spin-1/2 baryons (strongly interacting particles) has eight members: n^0, p^+, $\Sigma^\pm$, Σ^0, Ξ^-, Ξ^0, and Λ^0. Their strangeness quantum numbers are: S = 0: n^0, p^+; S = –1: $\Sigma^\pm$, Σ^0, and Λ^0; and S = –2: $\Xi^{0,-}$. If the positions of these eight particles are given in charge-strangeness space, a remarkable pattern emerges:



There are two particles at the center, each with zero charge and zero strangeness; they are the $\sum^0$ and the $\Lambda^{0.}$ (They have different rest masses).

They studied the structure of other families. A particularly important set of particles consists of all baryons with spin 3/2. At the time, there were nine known particles in this category: Δ^0, $\Delta^{\pm 1}$, Δ^{+2}, $\sum^{*0}$, $\sum^{*\pm 1}$, Ξ^0, and Ξ^{-1}. They have the following pattern in charge-strangeness space:



The symmetry pattern of the family of spin-3/2 baryons, shown by the known nine objects was sufficiently compelling for Gell-Mann, in 1962, to suggest that a tenth member of the family should exist. Furthermore, if the symmetry has a physical basis, the tenth member should have spin-3/2, charge –1, strangeness –3, and its mass should be about 150MeV greater than the mass of the Ξ^0 particle. Two years after this suggestion, the tenth member of the family was identified in high energy particle collisions; it

decayed via weak interactions, and possessed the predicted properties. This could not have been by chance. The discovery of the Ω^- particle was crucial in helping to establish the concept of the Gell-Mann – Ne'eman symmetry model.

In addition to the symmetries of baryons, grouped by their spins, the model was used to obtain symmetries of mesons, also grouped by their spins.

13.3 The special unitary group SU(3) and particle structure

Several years before the work of Gell-Mann and Ne'eman, Sakata had attempted to build-up the known particles from {neutron- proton- lambda^0} triplets. The lambda particle was required to "carry the strangeness". Although the model was shown not to be valid, Ikeda et al. (1959) introduced an important mathematical analysis of the three-state system that involved the group SU(3). The notion that an underlying group structure of elementary particles might exist was popular in the early 1960's. (Special Unitary Groups were used by J. P. Elliott in the late1950's to describe symmetry properties of light nuclei).

The problem facing Particle Physicists, at the time, was to find the appropriate group and its fundamental representation, and to construct higher-dimensional representations that would account for the wide variety of symmetries observed in charge-strangeness space. We have seen that the charge of a particle can be written in terms of its isospin, a concept that has its origin in the charge-independence of the nucleon-nucleon force.

When appropriate, we shall discuss the symmetry properties of particles in *isospin-strangeness space.*

Previously, we discussed the properties of the Lie group SU(2). It is a group characterized by its three generators, the Pauli spin matrices. Two-state systems, such as the electron with its quantized spin-up and spin-down, and the isospin states of nucleons and nuclei, can be treated quantitatively using this group. The symmetries of nucleon and meson families discovered by Gell-Mann and Ne'eman, implied an underlying structure of nucleons and mesons. It could not be a structure simply associated with a two-state system because the observed particles were endowed not only with positive, negative, and zero charge but also with strangeness. A three-state system was therefore considered necessary, at the very least; the most promising candidate was the group SU(3). We shall discuss the infinitesimal form of this group, and we shall find a suitable set of generators.

13.3.1 The algebra of SU(3)

The group of special unitary transformations in a 3-dimensional complex space is defined as

$$SU(3) \equiv \{\mathbf{U}_{3\times3} : \mathbf{U}\mathbf{U}^{\dagger} = \mathbf{I}, \det\mathbf{U} = +1, u_{ij} \quad \mathbf{C}\}.$$

The infinitesimal form of SU(3) is

$$SU(3)_{inf} = \mathbf{I} + i\delta\alpha_j\lambda_j/2 \ , \ j = 1 \text{ to } 8.$$

(There are $n^2 - 1 = 8$ generators).

The quantities $\delta\alpha_j$ are real and infinitesimal, and the 3×3 matrices λ_j are the linearly independent generators of the group. The repeated index, j, means that a sum over j is taken.

The defining properties of the group restrict the form of the generators. For example, the unitary condition is

$$\mathbf{U}\mathbf{U}^\dagger = (\mathbf{I} + i\delta\alpha_j\lambda_j/2)(\mathbf{I} - i\delta\alpha_j\lambda^\dagger_j/2)$$

$$= \mathbf{I} - i\delta\alpha_j\lambda_j^\dagger/2 + i\delta\alpha_j\lambda_j/2 \text{ to 1st-order,}$$

$$= \mathbf{I} \text{ if } \lambda_j = \lambda_j^\dagger.$$

The generators must be hermitian.

The determinantal condition is

$$\det = +1; \text{ and therefore } \mathrm{Tr}\lambda_j = 0.$$

The generators must be traceless.

The finite form of **U** is obtained by exponentiation:

$$\mathbf{U} = \exp\{i\alpha_j\lambda_j/2\}.$$

We can find a suitable set of 8 generators by extending the method used in our discussion of isospin, thus:

Let three fundamental states of the system be chosen in the simplest way, namely:

$$\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \text{ and } \mathbf{w} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

If we wish to transform $\mathbf{v} \rightarrow \mathbf{u}$, we can do so by defining the operator $\mathbf{A}_+$:

$$\mathbf{A}_+ \mathbf{v} = \mathbf{u}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

We can introduce other operators that transform the states in pairs, thus

$$\mathbf{A}_- = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{B}_+ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{B}_- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\mathbf{C}_+ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \mathbf{C}_- = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

These matrices are singular and non-hermitian. In the discussion of isospin and the group SU(2), the non-singular, traceless, hermitian matrices τ_1, and τ_2 are formed from the raising and lowering operators $\tau_\pm$ matrices by introducing the complex linear combinations

$$\tau_1 = \tau_+ + \tau_- = \sigma_1 \text{ and } \tau_2 = i(\tau_1 - \tau_2) = \sigma_2.$$

The generators of SU(3) are formed from the operators $\mathbf{A}_\pm$, $\mathbf{B}_\pm$, $\mathbf{C}_\pm$ by constructing complex linear combinations. For example:

the isospin operator $\tau_1 = \sigma_1 = \tau_+ + \tau_-$, a generator of SU(2) becomes

$$\begin{pmatrix} & & 0 \\ \sigma_1 & & 0 \\ 0 & 0 & 0 \end{pmatrix} = \mathbf{A}_+ + \mathbf{A}_- \equiv \lambda_1, \text{ a generator of SU(3).}$$

Continuing in this way, we obtain

$$\mathbf{A}_+ = \lambda_1/2 + i\lambda_2/2 ,$$

where

$$\lambda_2 = \begin{pmatrix} & & 0 \\ & \sigma_2 & \\ 0 & 0 & 0 \end{pmatrix},$$

and

$$\mathbf{C}_+ + \mathbf{C}_- = \lambda_4, \qquad \mathbf{C}_+ - \mathbf{C}_- = -i\lambda_5,$$

$$\mathbf{B}_+ + \mathbf{B}_- = \lambda_6 \text{ and } \mathbf{B}_+ - \mathbf{B}_- = i\lambda_7 .$$

The remaining generators, λ_3 and λ_8 are traceless, diagonal, 3×3 matrices:

$$\lambda_3 = \begin{pmatrix} & & 0 \\ & \sigma_3 & \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_8 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

The set of matrices $\{\lambda_1, \ldots.\lambda_8\}$ are called the Gell-Mann matrices, introduced in 1961. They are normalized so that

$$\mathrm{Tr}(\lambda_j\lambda_k) = 2\delta_{jk}.$$

The normalized form of λ_8 is therefore

$$\lambda_8 = (1/\sqrt{3}) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

If we put $\mathbf{F}_i = \lambda_i/2$. we find

$$\mathbf{A}_{\pm} = \mathbf{F}_1 \pm i\mathbf{F}_2 ,$$

$$\mathbf{B}_{\pm} = \mathbf{F}_6 \pm i\mathbf{F}_7,$$

and

$$\mathbf{C}_{\pm} = \mathbf{F}_4 \mp i\mathbf{F}_5 .$$

Let $\mathbf{A}_3 = \mathbf{F}_3$, $\mathbf{B}_3 = -\mathbf{F}_3/2 + (\sqrt{3}/4)\mathbf{F}_8$, and $\mathbf{C}_3 = (-1/2)\mathbf{F}_3 - (\sqrt{3}/4)\mathbf{F}_8$., so that

$$\mathbf{A}_3 + \mathbf{B}_3 + \mathbf{C}_3 = \mathbf{0}.$$

The last condition means that only eight of the nine operators are independent.

The generators of the group are readily shown to obey the Lie commutation relations

$$[\mathbf{F}_i, \mathbf{F}_j] = if_{ijk}\mathbf{F}_k , \; i,j,k = 1 \text{ to } 8.$$

where the quantities f_{ijk} are the non-zero structure constants of the group; they are found to obey

$$f_{ijk} = -f_{jik},$$

and the Jacobi identity.

The commutation relations $[\mathbf{F}_i, \mathbf{F}_j]$ can be written in terms of the operators $\mathbf{A}_{\pm}$, ...Some typical results are

$$[\mathbf{A}_+, \mathbf{A}_-] = 2\mathbf{A}_3, \; [\mathbf{A}_+, \mathbf{A}_3] = -\mathbf{A}_+, \; [\mathbf{A}_-, \mathbf{A}_3] = +\mathbf{A}_-,$$

$$[\mathbf{A}_3, \mathbf{B}_3] = 0, \quad [\mathbf{A}_3, \mathbf{C}_3] = 0, \quad [\mathbf{B}_3, \mathbf{C}_3] = 0$$

$$[\mathbf{B}_+, \mathbf{B}_-] = 2\mathbf{B}_3, \; [\mathbf{B}_+, \mathbf{B}_3] = -\mathbf{B}_-, \; [\mathbf{B}_-, \mathbf{B}_3] = +\mathbf{B}_-, \text{ etc.}$$

The two diagonal operators commute:

$$[\mathbf{F}_3, \mathbf{F}_8] = 0 \ .$$

Now, $\mathbf{F}_1$, $\mathbf{F}_2$, and $\mathbf{F}_3$ contain the 2×2 isospin operators (Pauli matrices), each with zeros in the third row and column; they obey the commutation relations of isospin. We therefore make the identifications

$$\mathbf{F}_1 = \mathbf{I}_1, \mathbf{F}_2 = \mathbf{I}_2, \text{ and } \mathbf{F}_3 = \mathbf{I}_3$$

where the $\mathbf{I}_j$'s are the components of the isospin.

Particles that experience the strong nuclear interaction are called *hadrons*; they are separated into two sets: the *baryons*, with half-integer spins, and the mesons with zero or integer spins. Particles that do not experience the strong interaction are called *leptons*. In order to quantify the difference between baryons and leptons, it has been found necessary to introduce the baryon number B = +1 to denote a baryon, B = –1 to denote an anti-baryon and B = 0 for all other particles. Leptons are characterized by the lepton number L = +1, anti-leptons are assigned L = –1, and all other particles are assigned L = 0. It is a present-day fact, based upon numerous observations, that the total baryon and lepton number in any interaction is conserved. For example, in the decay of the free neutron we find

$$n^0 = p^+ + e^- + \bar{\nu}_e{}^0$$

$$B = +1 = +1 + 0 + 0$$

$$L = 0 = 0 + 1 + (-1) \ .$$

The fundamental symmetries in Nature responsible for these conservation laws are not known at this time. These conservation laws may, in all likelihood, be broken.

In discussing the patterns of baryon families in charge-strangeness space, we wish to incorporate the fact that we are dealing with *baryons* that interact via the strong nuclear force in which *isospin* and *strangeness* are conserved. We therefore choose to describe their patterns in isospin-hypercharge space, where the hypercharge Y is defined to include both the strangeness and the baryon attribute of the particle in an additive way:

$$Y = B + S.$$

The diagonal operator $\mathbf{F}_8$ is therefore assumed to be directly associated with the hypercharge operator,

$$\mathbf{F}_8 = (\sqrt{3}/2)\mathbf{Y}.$$

Because $\mathbf{I}_3$ and $\mathbf{Y}$ commute, states can be chosen that are simultaneous eigenstates of the operators $\mathbf{F}_3$ and $\mathbf{F}_8$. Since no other SU(3) operators commute with $\mathbf{I}_3$ and $\mathbf{Y}$, no other additive quantum numbers are associated with the SU(3) symmetry. The operators $\mathbf{F}_4$,...$\mathbf{F}_8$ are considered to be new constants-of-the-motion of the strong interaction hamiltonian.

13.4 Irreducible representations of SU(3)

In an earlier discussion of the irreducible representations of SU(2), we found that the commutation relations of the generators of the group were satisfied not only by the fundamental 2×2 matrices but also by

matrices of higher dimension [(2J + 1) (2J + 1)], where J can have the values 1/2, 1, 3/2, 2, ...The J-values correspond to the spin of the particle whose state is given by a spinor (a column vector with special transformation properties). In the 2×2 representation, both *covariant* and *contravariant* spinors are allowed:

i) covariant spinors (with lower indices) are written as 2-component columns that transform under **U** SU(2) as

$$\alpha_i' = \mathbf{U}_i^j \alpha_j,$$

where

$$\alpha = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix},$$

and

ii) contravariant spinors (with upper indices) are written as 2-component rows that transform as:

$$\bar{\beta}^{j\prime} = \bar{\beta}^i \mathbf{U}_i^{j\,\dagger},$$

where

$$\bar{\beta} = (\bar{b}^1, \bar{b}^2).$$

The co- and contra-variant spinors are transformed with the aid of the anti-symmetric tensors ε_{ij} and ε^{ij}. For example,

$$\beta_i = \varepsilon_{ij} \bar{\beta}^j$$

transforms as a covariant spinor with the form

$$\beta_i = \begin{pmatrix} \bar{b}^2 \\ -\bar{b}^1 \end{pmatrix}.$$

The higher-dimensional representations are built up from the fundamental form by taking tensor products of the fundamental spinors α_i , $\bar{\beta}^j$, or β_i and by symmetrizing and anti-symmetrizing the result. We state, without proof, the theorem that is used in this method:

when a tensor product of spinors has been broken down into its symmetric and anti-symmetric parts, it has been decomposed into irreducible representations of the SU(n). (See Wigner's standard work for the original discussion of the method, and de Swart in Rev. Mod. Phys. **35**, (1963) for a detailed discussion of tensor analysis in the study of the irreps of SU(n))

As an example, we write the tensor product of two covariant spinors μ_i and ν_j in the following way

$$\mu_i \quad \nu_j = \mu_i\nu_j = (\mu_i\nu_j + \mu_j\nu_i)/2 \; + \; (\mu_i\nu_j - \mu_j\nu_i)/2$$

There are four elements associated with the product (i,j can have values 1 and 2).

The symmetric part of the product has three independent elements, and transforms as an object that has spin J=1. (There are 2J + 1 members of the symmetric set). The anti-symmetric part has one element, and therefore transforms as an object with spin J = 0. This result is familiar in the theory of angular momentum in Quantum Mechanics. The explicit forms of the four elements are:

$$J = 1 \quad \begin{cases} J_3 = +1: \ \mu_1\nu_1 \\ J_3 = 0 \ : (1/\sqrt{2})(\mu_1\nu_2 + \mu_2\nu_1) \\ J_3 = -1 \ : \mu_2\nu_1 \end{cases}$$

and

$$J = 0 \qquad J_3 = 0 : (1/\sqrt{2})(\mu_1\nu_2 - \mu_2\nu_1) \, .$$

Higher-dimensional representations are built up from the tensor products of covariant and contravariant 3-spinors, α and $\bar{\beta}$ respectively. The products are then written in terms of their symmetric and anti-symmetric parts in order to obtain the irreducible representations. For example, the product $\alpha_i\bar{\beta}^j$, i,j = 1,2,3, can be written

$$\alpha_i\bar{\beta}^j = (\alpha_i\bar{\beta}^j - (1/3)\delta_i^j\alpha_k\bar{\beta}^k) + (1/3)\delta_i^j\alpha_k\bar{\beta}^k \, ,$$

in which the trace has been separated out. The trace is a zero-rank tensor with a *single* component. The other tensor is a traceless, symmetric tensor with *eight* independent components. The decomposition is written symbolically as:

$$\mathbf{3} \quad \bar{\mathbf{3}} = \mathbf{8} \quad \mathbf{1}.$$

We can form the tensor product of two covariant 3-spinors, $\mu_i\nu_j$ as follows:

$$\mu_i\nu_j = (1/2)(\mu_i\nu_j + \mu_j\nu_i) + (1/2)(\mu_i\nu_j - \mu_j\nu_i), \ i,j = 1,2,3.$$

Symbolically, we have

$$\mathbf{3} \quad \bar{\mathbf{3}} = \mathbf{6} \quad \mathbf{3} \, ,$$

in which the symmetric tensor has six components and the anti-symmetric tensor has three components.

Other tensor products that will be of interest are

$$\mathbf{3} \quad \mathbf{3} \quad \mathbf{3} = \mathbf{10} \quad \mathbf{8} \quad \mathbf{8} \quad \mathbf{1}\,,$$

and

$$\mathbf{8} \quad \mathbf{8} = \mathbf{27} \quad \mathbf{10} \quad \overline{\mathbf{10}} \quad \mathbf{8} \quad \mathbf{8}' \quad \mathbf{1}\,.$$

The appearance of the octet "8" in the $\mathbf{3} \quad \overline{\mathbf{3}}$ decomposition (recall the observed octet of spin-1/2 baryons), and the decuplet "10" in the triple product $\mathbf{3} \quad \mathbf{3} \quad \mathbf{3}$ decomposition (recall the observed decuplet of spin-3/2 baryons), was of prime importance in the development of the group theory of "elementary" particles.

13.4.1 Weight diagrams

Two of the Gell-Mann matrices, λ_3 and λ_8, are diagonal. We can write the eigenvalue equations:

$$\lambda_3\mathbf{u} = \alpha_u\mathbf{u},\ \lambda_3\mathbf{v} = \alpha_v\mathbf{v},\ \text{and}\ \lambda_3\mathbf{w} = \alpha_w\mathbf{w},$$

and

$$\lambda_8\mathbf{u} = \beta_u\mathbf{u},\ \lambda_8\mathbf{v} = \beta_v\mathbf{v},\ \text{and}\ \lambda_8\mathbf{w} = \beta_w\mathbf{w}\,,$$

where α_i and β_i are the eigenvalues.

Let a and b be normalization factors associated with the operators λ_3 and λ_8, repectively, so that

$$\lambda_3{}^N = \begin{pmatrix} a & 0 & 0 \\ 0 & -a & 0 \\ 0 & 0 & 0 \end{pmatrix},\ \text{and}\ \lambda_8{}^N = \begin{pmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & -2b \end{pmatrix}.$$

If

$\mathbf{u} = [1, 0, 0]$, $\mathbf{v} = [0, 1, 0]$, and $\mathbf{w} = [0, 0, 1]$ (columns), we find

$$\lambda_3{}^N\mathbf{u} = a\mathbf{u}\,, \quad \lambda_8{}^N\mathbf{u} = b\mathbf{u},$$

$$\lambda_3{}^N\mathbf{v} = -a\mathbf{v}\,, \quad \lambda_8{}^N\mathbf{v} = b\mathbf{v}\,,$$

and

$$\lambda_3{}^N\mathbf{w} = 0\mathbf{w}\,, \quad \lambda_8{}^N\mathbf{w} = -2b\mathbf{w}.$$

The *weight vectors* are formed from the pairs of eigenvalues:

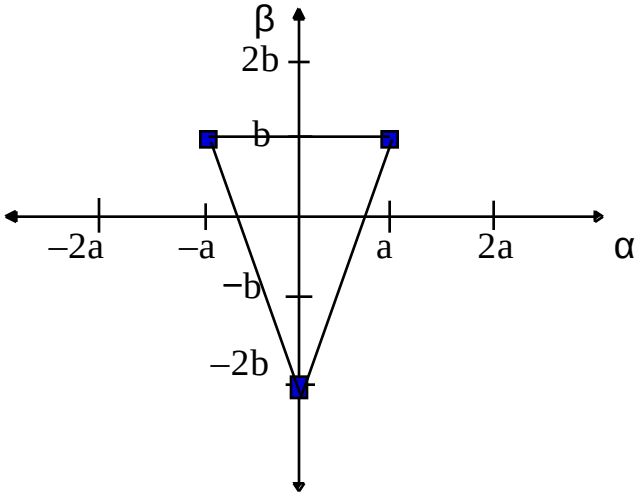
$$[\alpha_u, \beta_u] = [a, b],$$

$$[\alpha_v, \beta_v] = [-a, b],$$

and

$$[\alpha_w, \beta_w] = [0, -2b].$$

A weight diagram is obtained by plotting these vectors in the α–β space, thus:



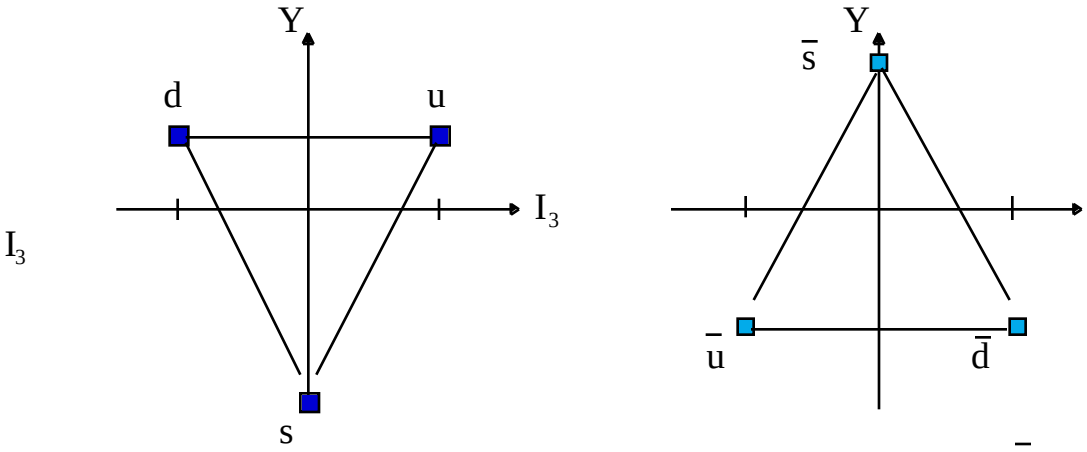
This weight diagram for the fundamental "3" representation of SU(3) was well-known to Mathematicians at the time of the first use of SU(3) symmetry in Particle Physics. It was to play a key role in the development of the quark model.

13.5 The 3-quark model of matter

Although the octet and decuplet patterns of hadrons of a given spin and parity emerge as irreducible representations of the group SU(3), major problems remained that resulted in a great deal of scepticism concerning the validity of the SU(3) model of fundamental particles. The most pressing problem was: why are there no known particles associated with the fundamental triplets **3**, $\bar{\mathbf{3}}$ of SU(3) that exhibit the symmetry of the weight diagram discussed in the last section? In 1964, Gell-Mann, and independently, Zweig, proposed that three fundamental entities do exist that correspond to the base states of SU(3), and that they form bound states of the hadrons. That such entities have not been observed in the free state is related to their enormous binding energy. The three entities were called *quarks* by Gell-Mann, and *aces* by Zweig. The Gell-Mann term has survived. The anti-quarks are associated with the conjugate $\bar{\mathbf{3}}$ representation. The three quarks, denoted by u, d, and s (u and d for the up- and down-isospin states, and s for strangeness) have highly unusual properties; they are

Label	B	Y	I	I_3	$Q = I_3 + Y/2$	$S = Y - B$
u	1/3	1/3	1/2	+1/2	+2/3	0
d	1/3	1/3	1/2	–1/2	–1/3	0
s	1/3	–2/3	0	0	–1/3	–1
$\bar{s}$	–1/3	2/3	0	0	+1/3	+1
$\bar{d}$	–1/3	–1/3	1/2	+1/2	+1/3	0
$\bar{u}$	–1/3	–1/3	1/2	–1/2	–2/3	0

The quarks occupy the following positions in I_3 - Y space



These diagrams have the same relative forms as the **3** and $\bar{\mathbf{3}}$ weight diagrams of SU(3).

The baryons are made up of quark triplets, and the mesons are made up of the simplest possible structures, namely quark–anti-quark pairs. The covariant and contravariant 3-spinors introduced in the previous section are now given physical significance:

μ = [u, d, s], a covariant column 3-spinor,

and

$\bar{\nu}$ = ($\bar{u}$, $\bar{d}$, $\bar{s}$), a contravariant row 3-spinor.

where u = [1, 0, 0], d = [0, 1, 0], and s = [0, 0, 1] represent the unitary symmetry part of the total wavefunctions of the three quarks.

The formal operators $\mathbf{A}_{\pm}$, $\mathbf{B}_{\pm}$, and $\mathbf{C}_{\pm}$, introduced in section **13.3.1,** are now viewed as operators that transform one *flavor* (type)of quark into another flavor (they are *shift* operators):

$$\mathbf{A}_{\pm} \equiv \mathbf{I}_{\pm}(I_3) \quad \rightarrow I_3 \pm 1 \ ,$$

$$\mathbf{B}_{\pm} \equiv \mathbf{U}_{\pm}(U_3) \rightarrow U_3 \pm 1, \text{ called the U-spin operator,}$$

and

$$\mathbf{C}_{\pm} \equiv \mathbf{V}_{\pm}(V_3) \rightarrow V_3 \pm 1, \text{ called the V-spin operator.}$$

Explicitly, we have

$$\mathbf{I}_{+}(-1/2) \rightarrow 1/2 : d \rightarrow u$$

$$\mathbf{I}_{-}(+1/2) \rightarrow -1/2 : u \rightarrow d$$

$$\mathbf{U}_{+}(-1/2) \rightarrow 1/2 : s \rightarrow d$$

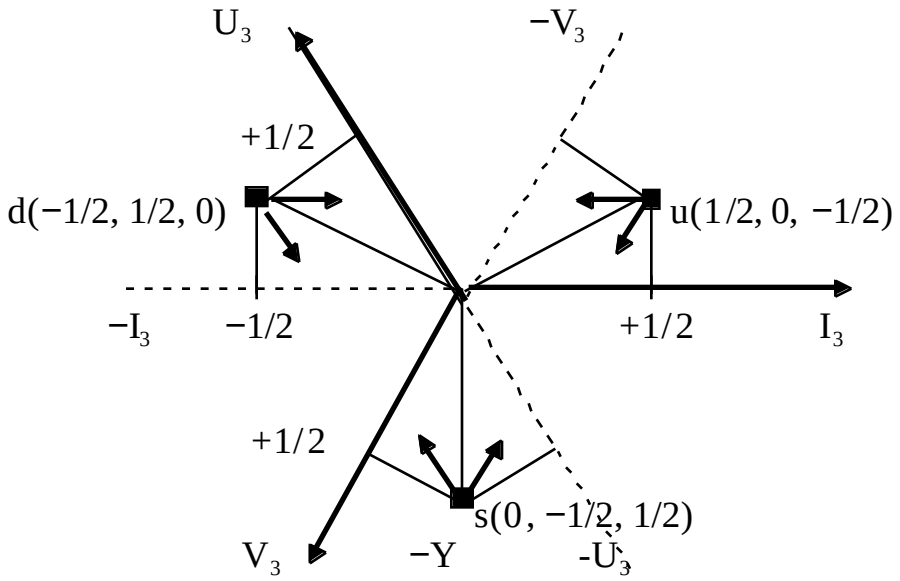
$$\mathbf{U}_{-}(+1/2) \rightarrow -1/2 : d \rightarrow s$$

$$\mathbf{V}_{+}(-1/2) \rightarrow 1/2 : u \rightarrow s$$

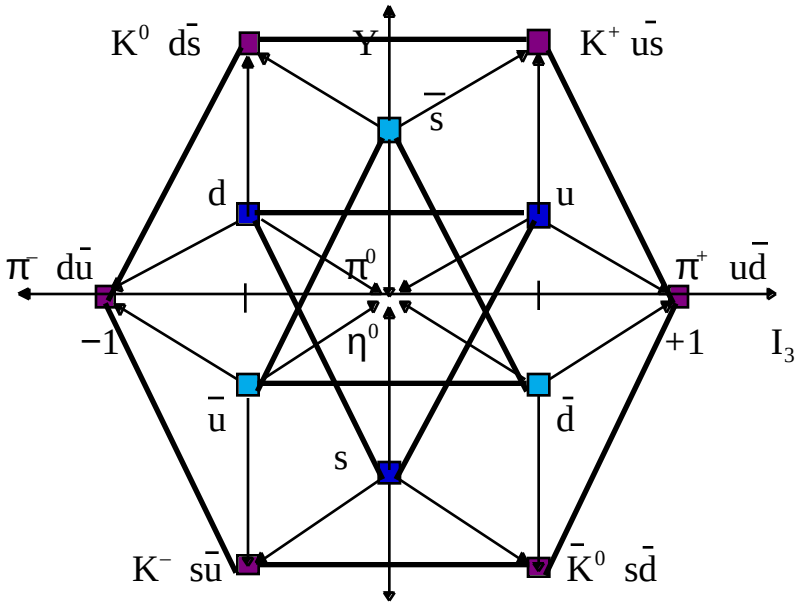
and

$$\mathbf{V}_{-}(+1/2) \rightarrow -1/2 : s \rightarrow u.$$

The quarks can be characterized by the three quantum numbers I_3, U_3, V_3. Their positions in the I_3-U_3-V_3 - space again show the underlying symmetry:



The members of the octet of mesons with $J^P = 0^-$ are formed from $q\bar{q}$-pairs that belong to the fundamental **3**, $\bar{\mathbf{3}}$ representation of the quarks. The π^0 and η^0 mesons are linear combinations of the $q\bar{q}$ states, thus



The nonet formed from the tensor product **3** $\bar{\mathbf{3}}$ is split into an octet that is even under the label exchange of two particles, and a singlet that is odd under label exchange:

$$\mathbf{3}\ \bar{\mathbf{3}} = \mathbf{8}\ \ \mathbf{1}$$

where the "1" is

$$\eta^{0\prime} = (1/\sqrt{3})(u\bar{u} + d\bar{d} + s\bar{s}),$$

and the two members of the octet at the center are:

$$\pi^0 = (1/\sqrt{2})(u\bar{u} - d\bar{d}) \text{ and } \eta^0 = (1/\sqrt{6})(u\bar{u} + d\bar{d} - 2s\bar{s}).$$

The action of I_- on π^+ is to transform it into a π^0. This operation has the following meaning in terms of I_- acting on the tensor product, u $\bar{d}$:

$$I_-(u\ \bar{d}) \equiv (I_-u)\ \bar{d} + u\ (I_-\bar{d}) \quad \text{(c.f. derivative rule)}$$
$$\downarrow \qquad\quad \downarrow \qquad\quad \downarrow$$
$$I_-(\pi^+) = d\ \bar{d} + u\ \bar{u}$$
$$\rightarrow \pi^0$$

Omitting the tensor product sign, normalizing the amplitudes, and choosing the phases in the generally accepted way, we have:

$$\pi^0 = (1/\sqrt{2})(u\bar{u} - d\bar{d}).$$

The singlet $\eta^{0'}$ is said to be *orthogonal* to π^0 and η^0 at the origin.

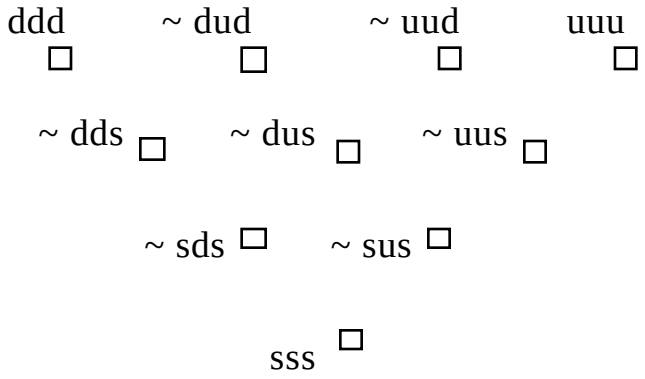
If the symmetry of the octet were exact, the eight members of the octet would have the same mass. This is not quite the case; the symmetry is broken by the difference in *effective* mass between the u- and d-quark (essentially the same effective masses: ~ 300 MeV/c^2) and the s-quark (effective mass ~ 500 MeV/c^2). (It should be noted that the effective masses of the quarks, derived from the mass differences of hadron-pairs, is not the same as the "current-quark" masses that appear in the fundamental theory. The discrepancy between the effective masses and the fundamental masses is not fully understood at this time).

The decomposition of **3** **3** **3** is

$$\mathbf{3}\ \mathbf{3}\ \mathbf{3} = (\mathbf{6}\ \mathbf{3})\ \mathbf{3}$$
$$= \mathbf{10}\ \mathbf{8}\ \mathbf{8'}\ \mathbf{1}$$

in which the states of the **10** are symmetric, the **1** is antisymmetric, and the **8**, **8′** states are of mixed symmetry. The decuplet that appears in this decomposition is associated with the observed decuplet of spin-3/2 baryons. In terms of the three fundamental quarks — u, d, and s, the make -up of

the individual members of the decuplet is shown schematically in the following diagram:



The precise make-up of each state, labelled by (Y, I, I_3,) is given in the following table:

$$
\begin{aligned}
(1, 3/2, +3/2) &= \mathbf{uuu}^{(++)} \\
(1, 3/2, +1/2) &= (1/\sqrt{3})(udu + duu + uud) \\
(1, 3/2, -1/2) &= (1/\sqrt{3})(ddu + udd + dud) \\
(1, 3/2, -3/2) &= \mathbf{ddd}^{(-)}
\end{aligned}
$$

$$
\begin{aligned}
(0, 1, +1) &= (1/\sqrt{3})(usu + suu + uus) \\
(0, 1, 0) &= (1/\sqrt{6})(uds + dsu + sud + dus + sdu + usd) \\
(0, 1, -1) &= (1/\sqrt{3})(dsd + sdd + dds)
\end{aligned}
$$

$$
\begin{aligned}
(-1, 1/2, +1/2) &= (1/\sqrt{3})(ssu + uss + sus) \\
(-1, 1/2, -1/2) &= (1/\sqrt{3})(ssd + dss + sds)
\end{aligned}
$$

$$
(-2, 0, 0) = \mathbf{sss}^{(-)}
$$

The general theory of the permutation group of n entities, and its representations, is outside the scope of this introduction. The use of the Young tableaux in obtaining the mixed symmetry states is treated in Hamermesh (1962).

The charges of the Δ^{++}, Δ^{-}, and Ω^{-} particles fix the fractional values of the quarks, namely

quark flavor	charge (in units of the electron charge)
u	+2/3
d	–1/3
s	–1/3

The charges of the anti-quarks are opposite in sign to these values.

Extensive reviews of the 3-quark model and its application to the physics of the low-energy part of the hadron spectrum can be found in Gasiorowicz (1966) and Gibson and Pollard (1976).

13.6 The need for a new quantum number: hidden color

Immediately after the introduction of the 3-quark model by Gell-Mann and Zweig, it was recognized that the model was not consistent with the extended Pauli principle when applied to bound states of three quarks. For example, the structure of the spin-3/2 Δ^+ state is such that, if each quark is assigned a spin $s_q = 1/2$, the three spins must be aligned $\uparrow\uparrow\uparrow$ to give a net spin of 3/2. (It is assumed that the relative orbital angular momentum of the quarks in the Δ^+ is zero (a symmetric s-state) — a reasonable assumption to make, as it corresponds to minimum kinetic energy, and therefore to a state of lowest total energy). The quarks are fermions, and therefore they must obey the generalized Pauli Principle; they cannot exist in a completely aligned spin state when they are in an s-state that is symmetric under particle (quark) exchange. The *unitary* spin component of the total wavefunction must be anti-symmetric. Greenberg (1964) proposed that a new degree of freedom must be assigned to the

quarks if the Pauli Principle is not to be violated. The new property was later called "color", a property with profound consequences. A quark with a certain flavor possesses color (red, blue, green, say) that corresponds to the triplet representation of another form of SU(3) — namely $SU(3)_C$, where the subscript C differentiates the group from that introduced by Gell-Mann and Zweig — the flavor group $SU(3)_F$. The anti-quarks (that possess anti-color) have a triplet representation in $SU(3)_C$ that is the conjugate representation (the $\bar{\mathbf{3}}$). Although the $SU(3)_F$ symmetry is known not to be exact, we have evidence that the *$SU(3)_C$ symmetry is an exact symmetry of Nature.* Baryons and mesons are found to be colorless; the color singlet of a baryon occurs in the decomposition

$$SU(3)_C = \mathbf{3} \quad \mathbf{3} \quad \mathbf{3} = \mathbf{10} + \mathbf{8} + \mathbf{8}' + \mathbf{1}\,.$$

The meson singlets consist of linear combinations of the form

$$\mathbf{1} = (R\bar{R} + B\bar{B} + G\bar{G})/\sqrt{3}\,.$$

Although the hadrons are colorless, certain observable quantities are directly related to the number of colors in the model. For example, the purely electromagnetic decay of the neutral pion, π^0, into two photons

$$\pi^0 = \gamma + \gamma,$$

has a lifetime that is found to be closely proportionl to the *square of the number of colors.* (Adler (1970) gives $\Gamma = \hbar/\tau = 1(\text{eV})\,(\text{number of colors})^2$

The measurements of the lifetime give a value of $\Gamma \sim 8$ eV, consistent with $N_{colors} = 3$. Since these early measurements, refined experiments have

demonstrated that there are three, and only three, colors associated with the quarks.

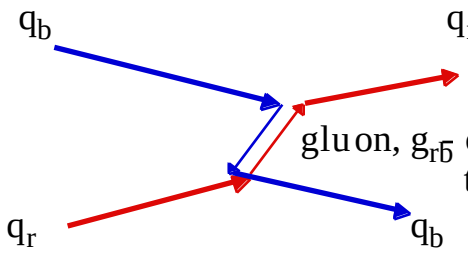
In studies of electron-positron interactions in the GeV-region, the ratio of cross sections:

$$R = \sigma(e^+e^- \rightarrow \text{hadrons})/\sigma(e^+e^- \rightarrow \mu^+\mu^-)$$

is found to depend linearly on the number of colors. Good agreement between the theoretical model and the measured value of R, over a wide range of energy, is obtained for *three* colors.

The color attribute of the quarks has been responsible for the development of a theory of the strongly interacting particles, called quantum chromodynamics. It is a field theory in which the quarks are generators of a new type of field — the color field. The mediators of the field are called *gluons*; they possess *color*, the attribute of the source of the field. Consequently, they can interact with each other through the color field. This is a field quite unlike the electrodynamic field of classical electromagnetism, in which the field quanta do not carry the attribute of the source of the field, namely electric charge. The photons, therefore, do not interact with each other.

The gluons transform a quark of a particular color into a quark of a different color. For example, in the interaction between a red quark and a blue quark, the colors are exchanged. This requires that the exchanged gluon carry color and anti-color, as shown:



Three different colors permit nine different ways of coupling quarks and gluons. Three of these are red-red, blue-blue, and green-green that do not change the colors. A linear combination $\sim(R\rightarrow\bar{R} + B\rightarrow\bar{B} + G\rightarrow\bar{G})$ is symmetric in the color labels, and this combination is the singlet state of the group $SU(3)_C$. Eight gluons, each with two color indices, are therefore required in the 3-color theory of quarks.

13.7 More massive quarks

In 1974, the results of two independent experiments, one a study of the reaction $p + Be \rightarrow e^+ + e^-$ .. (Ting et al.) and the other a study of $e^+ + e^- \rightarrow$ hadrons ..(Richter et al) — showed the presence of a sharp resonance at a center-of-mass energy of 3.1 GeV. The lifetime of the resonant state was found to be $\sim 10^{-20}$ seconds — more than 10^3 seconds longer than expected for a state formed in the strong interaction. The resonant state is called the J/ψ. It was quickly realized that the state corresponds to the ground state of a new quark–anti-quark system, a bound state $c\bar{c}$, where c is a fourth, massive, quark endowed with one unit

of a new quantum number c, called "charm". The quantum numbers assigned to the c-quark are

$$J^P = 1/2^+,\ c = 1,\ Q/e = +2/3, \text{ and } B = 1/3.$$

Sound theoretical arguments for a fourth quark, carrying a new quantum number, had been put forward several years before the experimental observation of the J/ψ state. Since 1974, a complex set of states of the "charmonium" system has been observed, and their decay properties studied. Detailed comparisons have been made with sophisticated theoretical models of the system.

The inclusion of a charmed quark in the set of quarks means that the group $SU(4)_F$ must be used in place of the original Gell-Mann-Zweig group $SU(3)_F$. Although the $SU(4)_F$ symmetry is badly broken because the effective mass of the charmed quark is ~ 1.8 GeV/c^2, some useful applications have been made using the model. The fundamental representations are

[u, d, s, c], a covariant column spinor,

and

($\bar{u}$, $\bar{d}$, $\bar{s}$, $\bar{c}$), a contravariant row spinor.

The irreps are constructed in a way that is analogous to that used in $SU(3)_F$, namely, by finding the symmetric and anti-symmetric decompositions of the various tensor products. The most useful are:

$$\mathbf{4}\ \ \bar{\mathbf{4}} = \mathbf{15}\ \ \mathbf{1},$$

$$\mathbf{4}\ \ \mathbf{4} = \mathbf{10}\ \ \bar{\mathbf{6}},$$

$$\mathbf{4} \quad \mathbf{4} \quad \mathbf{4} = \mathbf{20}_{\text{sym}} \quad \mathbf{20}_{\text{mix}} \quad \mathbf{20'}_{\text{mix}} \quad \bar{\mathbf{4}}_{\text{anti}},$$

and

$$\mathbf{15} \quad \mathbf{15} = \mathbf{1} \quad \mathbf{15}_{\text{sym}} \quad \mathbf{15}_{\text{anti}} \quad \mathbf{20}_{\text{sym}} \quad \mathbf{45} \quad \bar{\mathbf{45}} \quad \mathbf{84}.$$

The "15" includes the non-charmed ($J^P = 0^-$) mesons and the following charmed mesons:

$$D^0 = c\bar{u},\ \bar{D}^0 = \bar{c}u,\ \text{mass} = 1863\text{MeV}/c^2,$$

$$D^+ = c\bar{d},\ D^- = \bar{c}d,\ \text{mass} = 1868\ \text{MeV}/c^2,$$

$$F^+ = c\bar{s},\ F^- = \bar{c}s,\ \text{mass} = 2.04\ \text{MeV}/c^2.$$

In order to discuss the baryons, it is necessary to include the quark spin, and therefore the group must be extended to $SU(8)_F$. Relatively few baryons have been studied in detail in this extended framework.

In 1977, well-defined resonant states were observed at energies of 9.4, 10.01, and 10.4 GeV, and were interpreted as bound states of another quark, the "bottom" quark, b, and its anti-partner, the $\bar{b}$. Mesons can be formed that include the b-quark, thus

$$B_{\bar{u}} = b\bar{u},\ B_d^{\ 0} = b\bar{d},\ B_s^{\ 0} = b\bar{s},\ \text{and}\ B_c = b\bar{c}\ .$$

The study of the weak decay modes of these states is currently fashionable.

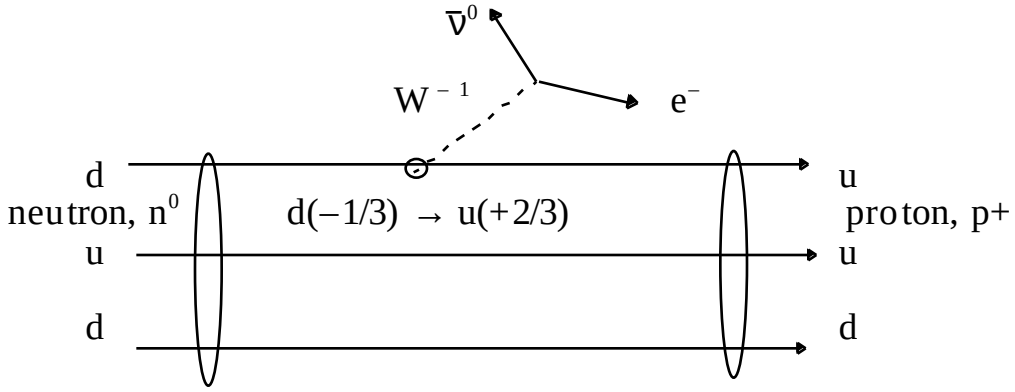
In 1994, definitive evidence was obtained for the existence of a sixth quark, called the "top" quark, t. It is a massive entity with a mass almost 200 times the mass of the proton!

We have seen that the quarks interact strongly via gluon exchange. They also take part in the weak interaction. In an earlier discussion of

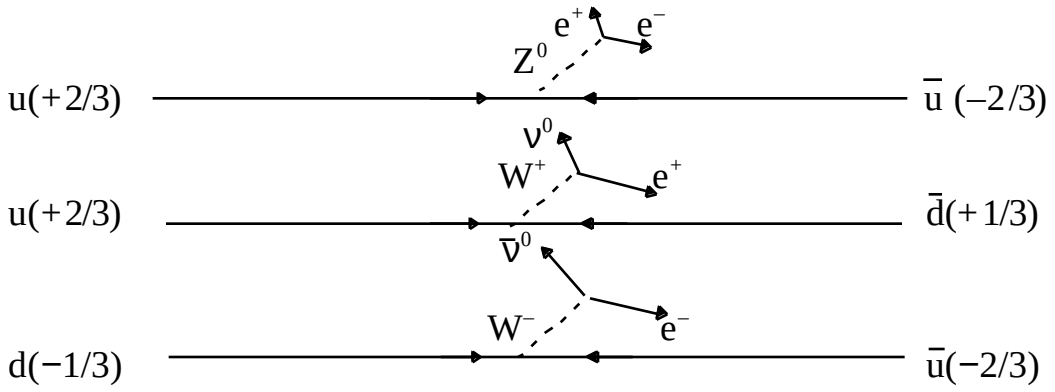
isospin, the group generators were introduced by considering the β-decay of the free neutron:

$$n^0 \rightarrow p^+ + e^- + \bar{\nu}^0 \ .$$

We now know that, at the microscopic level, this process involves the transformation of a d-quark into a u-quark, and the production of the carrier of the weak force, the massive W^- particle. The W^- boson (spin 1) decays instantly into an electron–anti-neutrino pair, as shown:



The carriers of the Weak Force, $W^{\pm}$, Z^0, were first identified in p-$\bar{p}$ collisions at high center-of-mass energy. The processes involve quark–anti-quark interactions, and the detection of the decay electrons and positrons.



The charge is conserved at each vertex.

The carriers have very large measured masses:

mass $W^{\pm} \sim 81$ GeV/c^2, and mass $Z^0 \sim 93$ GeV/c^2.

(Recall that the range of a force $\propto$ 1/(mass of carrier); the W and Z masses correspond to a very short range,~10^{-18} m, for the Weak Force).

Any quantitative discussion of current work using Group Theory to tackle Grand Unified Theories, requires a knowledge of Quantum Field Theory that is not expected of readers of this introductory book.

14

LIE GROUPS AND THE CONSERVATION LAWS OF THE PHYSICAL UNIVERSE

14.1 Poisson and Dirac Brackets

The *Poisson Bracket* of two differentiable functions

$$A(p_1, p_2, \ldots p_n, q_1, q_2, \ldots q_n)$$

and

$$B(p_1, p_2, \ldots p_n, q_1, q_2, \ldots q_n)$$

of two sets of variables $(p_1, p_2, ...p_n)$ and $(q_1, q_2, ...q_n)$ is defined as

$$\{A, B\} \equiv \sum_1^n (\partial A/\partial q_i)(\partial B/\partial p_i) - (\partial A/\partial p_i)(\partial B/\partial q_i) .$$

If $A \equiv \Omega(p_i, q_i)$, a dynamical variable, and

$B \equiv \mathbf{H}(p_i, q_i)$, the hamiltonian of a dynamical system,

where p_i is the (canonical) momentum and q_i is a (generalized) coordinate, then

$$\{\Omega, \mathbf{H}\} = \sum_1^n (\partial\Omega/\partial q_i)(\partial \mathbf{H}/\partial p_i) - (\partial\Omega/\partial p_i)(\partial \mathbf{H}/\partial q_i) .$$

(n is the"number of degrees of freedom" of the system).

Hamilton's equations are

$$\partial \mathbf{H}/\partial p_i = dq_i/dt \text{ and } \partial \mathbf{H}/\partial q_i = - dp_i/dt ,$$

and therefore

$$\{\Omega, \mathbf{H}\} = \sum_1^n (\partial\Omega/\partial q_i)(dq_i/dt) + (\partial\Omega/\partial q_i)(dp_i/dt) .$$

The total differential of $\Omega(p_i, q_i)$ is

$$d\Omega = \sum_1^n (\partial\Omega/\partial q_i)dq_i + (\partial\Omega/\partial p_i)dp_i.$$

and its time derivative is

$$(d\Omega/dt) = \sum_1^n (\partial\Omega/\partial q_i)(dq_i/dt) + (\partial\Omega/\partial p_i)(dp_i/dt)$$
$$= \{\Omega, \mathbf{H}\} = \dot{\Omega} .$$

If the Poisson Bracket is zero, the physical quantity Ω is a constant of the motion.

In Quantum Mechanics, the relation

$$(d\Omega/dt) = \{\Omega, \mathbf{H}\}$$

is replaced by

$$(d\Omega/dt) = -(i/\hbar))[\Omega, \mathbf{H}],$$

Heisenberg's equation of motion. It is the custom to refer to the *commutator* $[\Omega, \mathbf{H}]$ as the Dirac Bracket.

If the Dirac Bracket is zero, the quantum mechanical quantity Ω *is a constant of the motion*..

(Dirac proved that the classical Poisson Bracket $\{\Omega, \mathbf{H}\}$ can be identified with the Heisenberg commutator $-(i/\hbar)[\Omega, \mathbf{H}]$ by making a suitable choice of the *order* of the q's and p's in the Poisson Bracket).

14.2 Infinitesimal unitary transformations in Quantum Mechanics

The Lie form of an infinitesimal unitary transformation is

$$\mathbf{U} = \mathbf{I} + i\delta\alpha\mathbf{X}/\hbar\ ,$$

where $\delta\alpha$ ia real infinitesimal parameter, and $\mathbf{X}$ is an hermitian operator. (It is straightforward to show that this form of $\mathbf{U}$ is, indeed, unitary).

Let a dynamical operator Ω change under an infinitesimal unitary transformation:

$$\Omega \rightarrow \Omega' = \mathbf{U}\Omega\mathbf{U}^{-1}$$

$$= (\mathbf{I} + i\delta a\mathbf{X}/\hbar)\Omega(I - i\delta a\mathbf{X}/\hbar)$$

$$= \Omega - i\delta a\Omega\mathbf{X}/\hbar + i\delta a\mathbf{X}\Omega/\hbar \text{ to 1st-order}$$

$$= \Omega + i(\delta a\mathbf{X}\Omega - \Omega\delta a\mathbf{X})/\hbar$$

$$= \Omega + i(\mathbf{F}\Omega - \Omega\mathbf{F})/\hbar.$$

where

$$\mathbf{F} = \delta a\mathbf{X}.$$

The infinitesimal change in Ω is therefore

$$\delta\Omega = \Omega' - \Omega$$

$$= i[\mathbf{F}, \Omega]/\hbar$$

If we identify **F** with $-\mathbf{H}\delta t$ (the classical form for a purely temporal change in the system) then

$$\delta\Omega = i[-\mathbf{H}\delta t, \Omega]/\hbar,$$

or

$$-\delta\Omega = i[\mathbf{H}, \Omega]\delta t/\hbar\ ,$$

so that

$$-\delta\Omega/\delta t = i[\mathbf{H}, \Omega]/\hbar.$$

For a temporal change in the system, $\delta\Omega/\delta t = -\, d\Omega/dt$.

The fundamental Heisenberg equation of motion

$$d\Omega/dt = i[H,\Omega]/\hbar$$

is therefore deduced from the unitary infinitesimal transformation of the operator Ω.

This approach was taken by Schwinger in his formulation of Quantum Mechanics.

$|\mathbf{F}| = \mathbf{H}\delta t$ is directly related to the generator, **X**, of a Quantum Mechanical infinitesimal transformation, and therefore we can associate with every symmetry transformation of the system an hermitian operator **F** that is a constant of the motion - its eigenvalues do not change with time. This is an example of *Noether's Theorem*:

A conservation law is associated with every symmetry of the equations of motion. If the equations of motion are unchanged by the transformations of a Group then a property of the system will remain constant as the system evolves with time. As a well-known example, if the equations of motion of an object are invariant under translations in space, the linear momentum of the system is conserved.

15

BIBLIOGRAPHY

The following books are typical of those that are suitable for Undergraduates:

Armstrong, M. A., *Groups and Symmetry*, Springer-Verlag, New York, 1988.

Burns, Gerald, *Introduction to Group Theory*, Academic Press, New York, 1977.

Fritzsch, Harald, *Quarks: the Stuff of Matter*, Basic Books, New York, 1983.

Jones, H. F., *Groups, Representations and Physics*, Adam Hilger, Bristol, 1990.

The following books are of a specialized nature; they are typical of what lies beyond the present introduction.

Carter, Roger; Segal, Graeme; and Macdonald, Ian, *Lectures on Lie Groups and Lie Algebras*, Cambridge University Press, Cambridge, 1995.

Commins, E. D., and Bucksbaum, P. H., *Weak Interactions of Leptons and Quarks*, Cambridge University Press, Cambridge, 1983

Dickson, L. H., *Linear Groups*, Dover, New York, 1960.

Eisenhart, L. P., *Continuous Groups of Transformations*, Dover, New York, 1961.

Elliott, J. P., and Dawber, P. G., *Symmetry in Physics, Vol. 1*,

Oxford University Press, New York, 1979.

Gell-Mann, Murray, and Ne'eman, Yuval, *The Eightfold Way*, Benjamin, New York, 1964.

Gibson, W. M., and Pollard, B. R., *Symmetry Principles in Elementary Particle Physics*, Cambridge University Press, Cambridge, 1976.

Hamermesh, Morton, *Group Theory and its Applications to Physical Problems*, Dover, New York, 1989.

Lichtenberg, D. B., *Unitary Symmetry and Elementary Particles*, Academic Press, New York, 1978.

Lipkin, Harry J., *Lie Groups for Pedestrians*, North-Holland, Amsterdam, 1966.

Lomont, J. S., *Appplications of Finite Groups*, Dover, New York, 1993.

Racah, G., *Group Theory and Spectroscopy*, Reprinted in CERN(61-68), 1961.

Wigner, E. P., *Group Theory and its Applications to the Quantum Mechanics of Atomic Spectra*, Academic Press, New York, 1959.