ADVANCES IN ASTROBIOLOGY AND BIOGEOPHYSICS

A. Hanslmeier

Habitability and Cosmic Catastrophes

Springer

Habitability and Cosmic Catastrophes

Advances in Astrobiology and Biogeophysics

springer.com

This series aims to report new developments in research and teaching in the interdisciplinary fields of astrobiology and biogeophysics. This encompasses all aspects of research into the origins of life – from the creation of matter to the emergence of complex life forms – and the study of both structure and evolution of planetary ecosystems under a given set of astro- and geophysical parameters. The methods considered can be of theoretical, computational, experimental and observational nature. Preference will be given to proposals where the manuscript puts particular emphasis on the overall readability in view of the broad spectrum of scientific backgrounds involved in astrobiology and biogeophysics.
The type of material considered for publication includes:

- Topical monographs
- Lectures on a new field, or presenting a new angle on a classical field
- Suitably edited research reports
- Compilations of selected papers from meetings that are devoted to specific topics

The timeliness of a manuscript is more important than its form which may be unfinished or tentative. Publication in this new series is thus intended as a service to the international scientific community in that the publisher, Springer-Verlag, offers global promotion and distribution of documents which otherwise have a restricted readership. Once published and copyrighted, they can be documented in the scientific literature.

Series Editors:

Dr. André Brack
Centre de Biophysique Moléculaire
CNRS, Rue Charles Sadron
45071 Orléans, Cedex 2, France
Brack@cnrs-orleans.fr

Dr. Christopher P. McKay
NASA Ames Research Center
Moffet Field
CA 94035, USA

Dr. Gerda Horneck
DLR, FF-ME
Radiation Biology
Linder Höhe
51147 Köln, Germany
Gerda.Horneck@dlr.de

Prof. Dr. H. Stan-Lotter
Institut für Genetik
und Allgemeine Biologie
Universität Salzburg
Hellbrunnerstr. 34
5020 Salzburg, Austria

Arnold Hanslmeier

Habitability and Cosmic Catastrophes

Springer

Dr. Arnold Hanslmeier
Univ. Graz
Inst. Astronomie
Universitätsplatz 5
8010 Graz
Austria
arnold.hanslmeier@uni-graz.at

ISBN: 978-3-540-76944-6 e-ISBN: 978-3-540-76945-3

DOI 10.1007/978-3-540-76945-3

Advances in Astrobiology and Biogeophysics ISSN: 1610-8957

Library of Congress Control Number: 2008936264

Cover design: eStudio Calamar S.L.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

The search for life in the universe is one of the most challenging topics of science. It is not a modern topic at all, since more than 100 years ago, it was speculated that on the Moon, there are oceans and seas; on Venus, there are swamps and also Mars is inhabitated. However, now we have the scientific background and the scientific tools to answer this question and it is also certain that the answer would have deep implications for our culture, philosophy, and religions. If we find that life has developed on other planets or satellites of giant planets, then this would be the final breakdown of our central position in the universe. But is life a widespread phenomenon? How vulnerable is it to changing conditions and even catastrophic events? These topics will be discussed in this book.

If life is in the extreme case a unique phenomenon found only on planet Earth, which seems to be highly unrealistic, then also it is important to discuss how it is adaptable to changing external conditions. Can we survive a cosmic catastrophe? How do these catastrophes change habitability? Which forms of life are more vulnerable?

It was mentioned that now science has made great progress to answer such questions. Let us give some examples.

In modern biology, in connection with organic chemistry, the origin of life is studied. Still some surprising discoveries have been made over the last decades such as life under extreme conditions, the extremophiles. The chemical elements that are necessary for life (e.g., carbon, oxygen) have been produced at the centers of stars by thermonuclear fusion reactions. Thus, life could not exist in the early universe because at that time only hydrogen and helium were present, both had been formed within the first three minutes after the Big Bang.

Recent astronomical observations have proven the existence of extrasolar planetary systems; more than 250 such planets have been detected. With new upcoming satellite missions it will be possible to detect directly Earth-sized extrasolar planetary systems and even to measure any sign of biologic activity there. Also, great discoveries were made in our solar system. Since the surface of Venus turned out to be too hot, only Mars remained a possible candidate; however, the discovery of satellites of Jupiter and Saturn that are covered by a thick ice crust and where there is a liquid water ocean beneath that crust extended the search for life in the solar system.

Thus, it has to be expected that the question whether there exists life on other planets or not will be answered within the next few decades.

The investigation of habitable conditions on a planet or moon of a planet has also to be seen in the context of possible catastrophes. That such catastrophes happened on Earth is now well established, e.g., 65 million years ago the dinosaurs and other species became extinct due to an impact of an asteroid. Are there other possible cosmic catastrophes that could be dangerous to the habitability on a planet? In this book, we will discuss about collisions, orbit instabilities, supernova explosions, intense stellar activity outbursts, activity in galaxies triggered by their supermassive black holes, and other topics which are really cosmic catastrophes and could destroy habitability on a planet.

Cosmic catastrophes lead to mass extinction and could make a planet inhabitable. However, they also provide a chance for new lifeforms to appear. On Earth, maybe without the extinction of the dinosaurs, mammals never would have had a chance to evolve and propagate so rapidly. Intense UV radiation leads to mutations, many of which become extinct but some of them provided big progress in the evolution.

The book is intended for students and teachers in science, biology, chemistry, astrophysics, and physics and it provides an overview of the above-mentioned topics; in the appendix, some details on organic chemistry and astrophysics are given. For the interested reader, about 250 recent papers and articles are cited. This will help the reader to penetrate deeper into special topics. In the beginning of the chapters, books and review articles that provide a deeper insight into the topics discussed are mentioned in the footnotes.

I would like to thank my family for their patience because I spent lots of nights on the manuscript. Many colleagues gave me advice on different topics and, especially, I greatly acknowledge the excellent cooperation with Springer, Mr. Ramon Khanna and Miss Hema Latha.

Graz/Austria, October 2008 *Arnold Hanslmeir*

Contents

List of Tables

Chapter 1
Habitable Zones

Habitable zones (HZ) around stars mean regions in which, on a hypothetical object, conditions favorable to life might be found. In this chapter, we discuss habitability on planets and satellites of planets. As we will show, habitability depends on different factors:

- Evolution, age, and activity of the mother star
- Atmosphere, magnetosphere, and distance of a planet from its mother star
- Stability of the planetary system
- Galactic environment of the planetary system

Before evaluating these parameters in more detail, we concentrate first on habitability. We will discuss the different attempts how these zones can be defined, how their extension and location depends on the central star and, of course, discuss in more detail the special conditions found in our solar system.

1.1 Definition of Habitability

1.1.1 Basic Considerations

Under the assumption that life evolves under favorable conditions, habitability on a planet is more or less a synonym for the existence of life on it. There is only one known example in the universe where we can investigate how life might have evolved – our Earth. Some basic principles of biochemistry and how life might have evolved will be discussed in Chap. 2.

1.1.1.1 The Host Star: The Energy Source

It is evident that life needs energy to maintain metabolism, cell growth, cell division, etc. This energy is mainly provided by the central star of a planetary system because planets themselves do not produce much energy. Only during the first few 100,000

A. Hanslmeier, *Habitability and Cosmic Catastrophes*, Advances in Astrobiology and Biogeophysics, DOI 10.1007/978-3-540-76945-3_1,

years, gravitational contraction and energy released by impacts played an important role on planets like the Earth. Though it is estimated that the geothermal energy in the upper 5 km of the Earth's crust is 40 millions times greater than the energy contained in the oil and gas reserves, without warming from the Sun, the Earth would be a cold frozen planet hostile to life. The internal heat was originally created by accretion when the Earth formed by contraction and then by radioactive decay of Uranium, Thorium, and potassium. The heat flow from the interior to the surface is only 1/20,000 as great as the energy received from the Sun.

Stars have different masses, ages, and evolutionary tracks (see also 3.1.4). HZ depend on the properties of the central star of a planetary system: mass, temperature, luminosity, age, and activity are important parameters of stars that strongly influence the location and extension of HZ. Stars are relatively easy to understand – at least during most of their lifetime. Let us briefly summarize some basic relations:

- Temperature: depends on the energy production rate (mainly by nuclear fusion). The more massive the stars, the higher the energy production rate and, as a consequence, the luminosity. The Sun with a surface temperature of about 6,000 K belongs to the cooler stars.
- Luminosity: the luminosity of a star depends on its effective temperature T_{eff} and its surface $4\pi R^2$:

$$L = 4\pi R^2 \sigma T_{\mathrm{eff}}^4 \tag{1.1}$$

  Stars with twice the radius of the Sun will be 4 times more luminous and stars with twice the surface temperature of the Sun will be 16 times more luminous than the Sun.
- Mass: the mass is the most important parameter for the evolution of a star. The higher the mass, the higher the temperature. For cool stars with masses $\leq 1\,\mathrm{M}_\odot$, where $\mathrm{M}_\odot$ is the mass of the Sun (2×10^{30} kg), the HZ will be closer to the star because the star generates less energy.
- Stellar evolution: stars become more luminous during their evolution. Therefore, the HZ tends to move outward. Generally, more massive stars evolve much faster than low massive stars.

1.1.1.2 Surface Temperature on Planets

The surface temperature on a planet is crucial for the existence of life. We can easily estimate the temperature on the surface of a planet by calculating the energy it receives from its host star and taking into account the distance between the planet and its host star.

Let $S(r)$ be the radiation flux emitted by a star at distance r from a planet and S_0 the surface flux, then

$$S(r) = S_0/(4\pi r^2) \tag{1.2}$$

A planet with albedo A absorbs

$$S(r)(1-A)\pi R^2 \tag{1.3}$$

According to the Stefan–Boltzmann law this equals:

$$\pi R^2(1-A)S(r) = 4\pi R^2 \sigma T^4 \tag{1.4}$$

If the planet has *internal heat sources* (For Earth this is negligible; due to the radioactive decay, only 0.06 $\mathrm{W/m^2}$ is produced. For the giant planets, there are considerable heat sources), then on the left hand side of equation (1.4), we must add the term $4\pi R^2 Q$.

Let us estimate the temperature expected on Earth from this equation. The Sun emits $3.86 \times 10^{26}\,\mathrm{W}$. The Sun–Earth distance is $150 \times 10^9\,\mathrm{m}$, $\sigma = 5.67 \times 10^{-8}\,\mathrm{Wm^{-2}K^{-4}}$, and the radius of the Earth is about $6.4 \times 10^6\,\mathrm{m}$. Inserting these values, we see that the Earth's radius does not enter and the expected surface temperature becomes

$$T = \left[\frac{1-A}{\sigma}\frac{3.86\times 10^{26}}{4\pi(1.5\times 10^{11})^2}\right]^{1/4} = 254.5\,\mathrm{K}$$

The albedo A of an object is the extent to which it diffusely reflects light from the Sun. The larger is the albedo, the lower will be the expected surface temperature because more radiation is reflected back. In this calculation, we have taken the value 0.3 for the albedo. This simple estimation does not take into account the atmosphere of a planet; on the Earth, due to the natural greenhouse warming, which is mainly caused by water vapor and carbon dioxide, the surface temperature is about 35° higher than the value from above; therefore, greenhouse gases are essential to maintaining the temperature on Earth but an excess of greenhouse gases can raise the temperature to a lethal level.

1.1.2 Definitions of Habitable Zones

To our knowledge, the life can only evolve in an environment that is quite similar to Earth. A habitable zone (HZ) can be defined as a region of space, where the conditions are favorable for life based on complex carbon compounds and on availability of fluid water, etc., as it can be found on Earth.

1.1.2.1 Circumstellar Habitable Zones

Let us consider stars of different types and ask the question whether there exists a HZ around them or not. In this context, often the term *circumstellar habitable zone* or ecosphere is used. Obviously, the Earth is located in the center of the HZ of the solar system. Let d denote the mean radius of the HZ in AU,[1] L_* the bolometric luminosity[2] of a star, and $L_\odot$ the bolometric luminosity of the Sun, then

[1] AU means astronomical unit, the mean Sun–Earth distance, $150\times 10^6\,\mathrm{km}$.

[2] Bolometric means a star's output over all wavelengths.

$$d = \sqrt{\frac{L_*}{L_\odot}} \tag{1.5}$$

so that for $L_* = L_\odot$, we obtain $d = 1$. The center of the HZ of a star with $0.25\,\mathrm{L}_\odot$ would be at a distance of 0.5 AU from the star. A star with $2\mathrm{L}_\odot$ would have the center of the HZ at a distance of 1.41 AU.

Another crude approximation can be made by estimating the solar flux. The flux varies with the square of the distance from the Sun and therefore,

$$F \sim 1/r^2 \tag{1.6}$$

Let us consider the orbital parameters of Venus and Mars. The semi-major axis of Venus is 0.7 AU, the semi-major axis of Mars is 1.4 AU. Then, the solar flux at the orbit of Venus is

$$F_{\mathrm{Venus}} \sim 1/(0.7)^2 \sim 2 \tag{1.7}$$

On Venus, the solar flux to be expected will be twice the flux on Earth; for Mars, we obtain

$$F_{\mathrm{Mars}} \sim 1/(1.4)^2 \sim 0.5 \tag{1.8}$$

The flux changes by a factor of 4 between the orbits of Venus and Mars. These basic considerations illustrate that the HZ even in a solar-like star will be restricted to a small circumstellar region.

Now let us consider the HZ for different stellar masses as shown in Fig. 1.1. The higher the mass, the greater the luminosity of a star and the more distant the region of the HZ to the mother star.

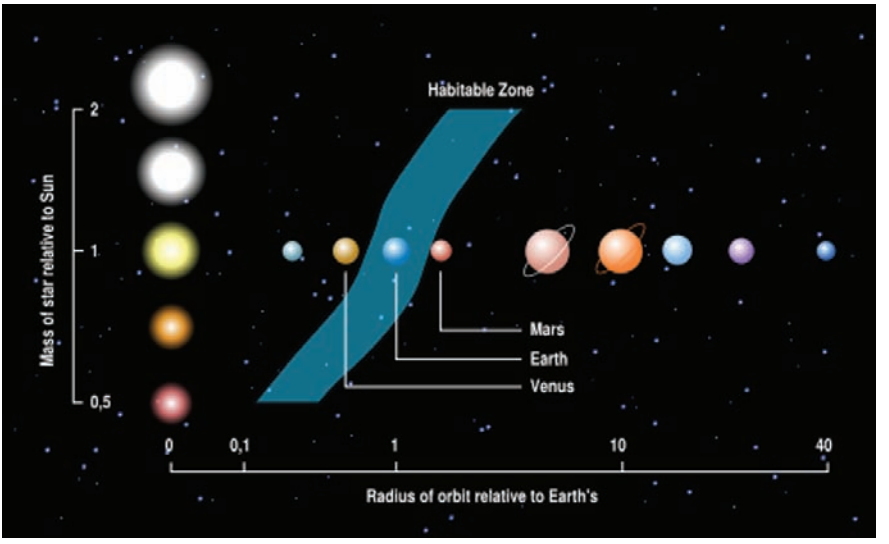

Fig. 1.1 The location of the habitable zone for different masses of the central stars (*left*). Note that the abscissa is in logarithmic units. The solar system (at 1 solar mass) is given at the center. Courtesy: Wikimedia Cosmos

1.1.2.2 Continuously Habitable Zones

So far, HZ were defined under the assumption that the luminosity of a central star remains constant. We know that the luminosity of the Sun changed from $0.75 L_\odot$ to its present value.

Habitable zones tend to move outward with time because main sequence stars become brighter. For life to evolve to higher forms, the planet must be continuously in the HZ that slowly progresses outward due to stellar evolution. This leads to the definition of a continuously habitable zone (also denoted as CHZ, not to be confused with circumstellar habitable zone!):

→The *continuously habitable zone* is the region in space where a planet remains habitable for some long time period $\tau_{\rm hab}$.

Since the time that was needed on Earth for intelligent life to evolve is about 4 Gy the same or a similar value is usually taken for $\tau_{\rm hab}$. Some other authors take smaller values such as $\tau_{\rm hab}$ =3 Gy or $\tau_{\rm hab}$ =1 Gy, the latter value holds only for the evolution of microbiological life [218].

One of the earliest attempts to define a HZ was made by Hart [95]. He gave also a formula for the inner and outer boundaries of the HZ:

$$r_o/r_i \sim [L(3.5)/L(1.0)]^{1/2} \tag{1.9}$$

Here, r_i denotes the inner boundary and r_o the outer boundary of a continuously habitable zone. $L(t)$ represents the luminosity after t billion years of a main sequence star (see Sect. 3.2.2) of mass M. His conclusion was that the ratio r_o/r_i becomes smaller for less massive stars converging to $r_o = r_i$ for stars with $M \sim 0.83 M_\odot$. That would implicate that low massive cool stars (of spectral type later than K1) have no continuously habitable zone.

Circumstellar habitable zones (CHZ) strongly depend on the UV radiation. On Earth, UV radiation between 200 and 300 nm is very damaging to biological systems. On the other hand, it must be taken into account that UV radiation on the primitive Earth was one of the most important energy sources for the synthesis of biochemical compounds and therefore essential to biogenesis. Buccino et al. [26] estimated the UV constraints on the HZ. Applying diverse criteria to those stellar systems whose central star has been observed by IUE,[3] they obtained that an Earth-like planet orbiting the stars *HD 216437*, *HD 114752*, *HD 89744*, τ *Boo*, and ρ *CrB* could be habitable for at least 3 Gyr. A moon orbiting ν *And c* would be also suitable for life, while in 59% of the sample (*51 Peg*, *16 Cyg B*, *HD 160691*, *HD 19994*, *70 Vir*, *14 Her*, *55 Cnc*, *47 UMa*, ε *Eri*, and *HD 365*), the traditional HZ would not be habitable.

[3] International Ultraviolet Explorer satellite

1.1.2.3 Noncircumstellar HZ

We should mention here that maybe these definitions are quite restrictive. If HZ are restricted to the presence of liquid water, then we must also include the interiors of giant planets and the ice-covered Galilean Moons of Jupiter, the Martian subsurface, and maybe even other places. Therefore, the search for life must not be restricted to the study of CHZ. But it seems that life requires much more than just liquid water.

Generally, for life to occur, there is also a need for an energy source. In the case of CHZ, this is provided by the host star of the system. In the cases of satellites of planets, this can be provided by other mechanisms such as tidal locking, which will be described in Sect. 1.1.4.

1.1.2.4 Galactic Habitable Zone

Also, the location of a planetary system in a galaxy is important for long-term habitability. This will be discussed in more detail in Sect. 1.1.6 of this chapter.

1.1.3 HZ and Planetary Atmospheres

We have already outlined that the natural greenhouse gases are essential for maintaining temperatures on Earth that are above the freezing point of water. The composition, extension, and thickness of planetary atmospheres are crucial parameters for habitability.

A more accurate definition of HZ including simple atmospheric models of planets was given by Kasting et al. [117]. They used a one-dimensional climate model on Earth-like planets with CO_2, H_2O, N_2 atmospheres and the above-mentioned condition of liquid water present on the planet's surface. These assumptions lead to an outer edge and an inner edge:

- Outer edge: this is defined by the distance of the planet from its central star at which the formation of CO_2 clouds starts. These clouds cool the planet's surface by increasing its albedo.
- Inner edge: because of the low distance of the planet to its central star, loss of water via photolysis occurs, water goes into the stratosphere, the H_2O molecules are split by UV radiation from the star into H_2 and O, and hydrogen, H_2, escapes from the atmosphere. This is also referred to as moist greenhouse effect.

If these constraints are applied to our solar system, then the limits for the HZ are 0.95 AU $<$ HZ $<$ 1.37 AU. However, these are still crude assumptions and the HZ could be much greater. Between these two limits climate stability is ensured by a simple feedback mechanism: let $< CO_2 >$ denote atmospheric CO_2 concentration and T_{surf} denote the surface temperature on a planet, then

$$< CO_2 > \sim 1/T_{surf} \tag{1.10}$$

It was also shown that the width of the HZ depends on the size of the planet. It increases for planets that are larger than Earth and also for planets that have higher N_2 partial pressures.

For our solar system, the continuously habitable zone can be estimated as

$$0.95 < \mathrm{CHZ} < 1.15\,\mathrm{AU} \tag{1.11}$$

Later it was shown by Forget, Pierrehumbert, and Ramond [63] and Mischna et al. [168] that for planets with high concentrations of CO_2 in their atmosphere, an additional warming effect can occur. This would extend the HZ beyond the limits found so far.

Franck et al. [64] defined their CHZ by surface temperature bounds in the interval 0°C and 100°C and a CO_2 partial pressure above 10^{-5} bar. They found for the solar system an inner boundary of 0.95 AU and an outer boundary of 1.2 AU.

1.1.4 Tidal Locking

A well-known example of tidal locking is the Earth–Moon system. From Earth, you can always see the same hemisphere of the Moon. We say that the Moon is *tidally locked*; it takes just as long to rotate around its own axis as it does to complete one revolution around the Earth. Also the term *synchronous rotation* is used. The Moon is 1:1 tidally locked.

Let us consider a smaller body B, which gets locked to a larger body A. The gravity of A produces a tidal force on B. This results in a deviation from its equilibrium configuration, i.e., B becomes slightly elongated and stretched along the axis oriented toward A and on the other hand, it is compressed in the two perpendicular directions → tidal bulges (see Fig. 1.2 and Fig. 1.3). The result is that B becomes an ellipsoid.

However, the material of B exerts resistance to this reshaping caused by the tidal force. There is also some delay: if the rotation period of B is shorter than its orbital period, the bulges are carried forward of the axis oriented toward A in the direction of rotation. The bulges lag behind it if the orbital period of B is shorter. This effect is known as bulge dragging.

The bulges are displaced from the $A - B$ axis and this results in a torque. Note that there are in fact two torques which react in opposite sense but are not exactly equal:

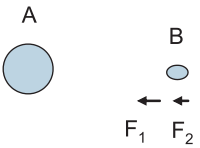

Fig. 1.2 Two bodies A and B attract each other, $M_A > M_B$. The gravitational attraction of A on B is larger on the surface oriented toward A (i.e., $F_1 > F_2$)

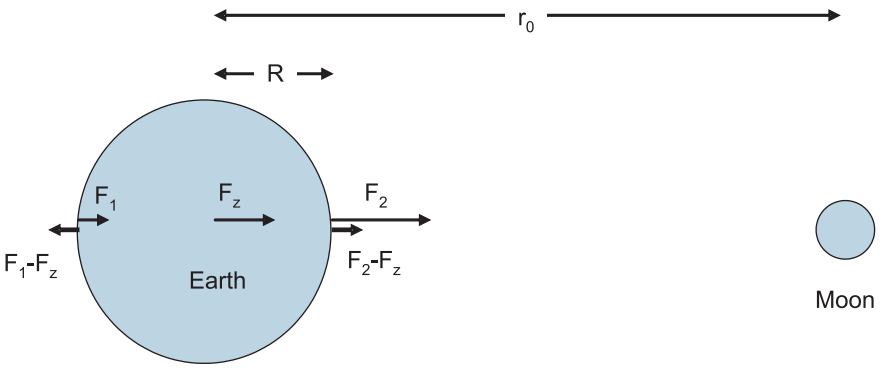

Fig. 1.3 Explanation of the two tidal bulges on Earth caused by the attraction of the Moon. F_2 denotes the force exerted by the moon on the face of the Earth toward it (i.e., the moon is in the highest point over the horizon there). F_z is the force that the moon exerts on the center of the Earth and F_1 is the force of the moon on the side opposite to it. Subtracting F_z from F_2 and F_1, we see that there is a resulting force that is pointing away from the Earth's surface, thus causing the tidal bulges

- Torque on the *A*-facing bulge acts to bring *B*'s rotation in line with its orbital period
- Torque on the side on *B* opposite to *A* acts in opposite sense
- The gravitational attraction on the *A*-facing side is slightly higher than on the opposite side
- → tidal locking

Conservation of angular momentum in the system results into an additional effect. Because *B* slows down it loses angular momentum and to conserve the total momentum, the momentum resulting from the orbital motion must increase, which means that the semi-major axis of the orbit of *B* increases.

$$L_{\text{orbit}} = a_{\text{orbit}} M_B v_{\text{orbit}} \tag{1.12}$$

Therefore, the scenario for object *B* is as follows:

Rotational slowdown → increase orbital diameter.

But there is another case: when *B* starts off rotating too slowly, tidal locking speeds up rotation and lowers the orbit.

Summarizing, tidal locking leads to two effects depending on the initial rotation of *B*:

- $P_{\text{rot,B}} < P_{\text{orbit,B}}$: rotation slowdown, orbit increases
- $P_{\text{rot,B}} > P_{\text{orbit,B}}$: rotation speedup, orbit decreases

So far, we have only considered the effects on the smaller body *B*. But the tidal locking effect of course has influence on the larger body *A*. The effects on *A* depend on the mass ratio M_B/M_A. In the case of the *Earth–Moon* system, this ratio is 1/81. This causes the Earth also to spin down gradually. Another example is the system

Pluto–Charon. These bodies are of similar size and mass and both are tidally locked to each other. Astronauts on Pluto would be able to observe Charon only from one hemisphere.

Another example is *Rotation–Orbit resonances*. This happens when the orbit is eccentric and the tidal effects are not too large. *Mercury* is in a 3:2 rotation–orbit resonance with the Sun. This means the planet makes three rotations after completing two orbits around the Sun.

The final configuration of such a tidally locked system is the one that occurs in the lowest energy. This means that the heavy side will face the planet when considering the case of a moon and a planet. It also becomes clear because of this configuration why planet-facing hemispheres are visibly different than the opposite faces. In the case of our Moon, this is evident. The hemisphere that is facing the Earth is quite different – there are the large maria which are impact basins filled with lava later. The maria are composed of basalt and this is heavier than the surrounding highland crust. However, this is only a strong simplification because the timescales are different: tidal locking occurred in a very short timescale (only some 10^3 y), the formation of the maria occurred on much larger timescales.

Let us consider the timescale for tidal locking (see [72] where further references on that topic can also be found):

$$t_{\text{lock}} \approx \frac{waIQ}{3GM_p^2 k_2 R^5} \tag{1.13}$$

w is the initial spin rate (revolutions per s), a the semi-major axis of the satellite, $I \sim 0.4 M_S R^2$ the moment of inertia of the satellite with mass M_S, Q the dissipation function of the satellite, G the gravitational constant, M_P the mass of the planet, k_2 the tidal Love number, and R the radius of the satellite. The tidal Love number is

$$k_2 \sim \frac{1.5}{1 + \frac{19\mu}{2\rho g R}} \tag{1.14}$$

where ρ is the density of the satellite, $g = GM_S/R^2$ the surface gravity of the satellite, and μ the rigidity of the satellite. This value depends on the material which forms the satellite and the two extremes are

$$\mu = 3 \times 10^{10}\,\text{Nm}^{-2} \rightarrow \text{rocky satellites}$$

$$\mu = 4 \times 10^{9}\,\text{Nm}^{-2} \rightarrow \text{icy satellites}$$

For the system Earth–Moon, the ratio $k_2/Q = 0.0011$ can be taken. As a first approximation, one can take $Q \sim 100$ and calculate k_2 according to the above given formula.

The formulae can be simplified with $k_2 << 1, Q = 100$ and assuming a revolution every 12 h in the initial nonlocked state. These values hold for asteroids in the solar system because they have orbital periods between 2 h and 2 d. Then we arrive at

$$t_{\text{lock}} \approx 6\,\frac{a^6 R\mu}{M_s M_p^2} \times 10^{10}\ \text{years} \tag{1.15}$$

Masses are in kg, distances in m, and μ in Nm^{-2}. Note that the tidal locking time strongly depends on the semi-major axis of the satellite. Let us consider the factor R/M_s in the formula which is $R/(4\pi/3R^3\rho_s) \sim 1/R^2$. This means that larger satellites will lock faster than smaller ones.

For our discussion about HZ, such resonances and tidal locks are extremely important. If, e.g., a planet always faces the same hemisphere to its central star, then this would have enormous effects for the climate and weather systems on it.

1.1.4.1 Hot Jupiters

Tidal locking has some other important consequences. Let us consider exoplanets orbiting close to their host stars. Then, their internal magnetic fields would be very weak because of the slow rotation due to tidal locking. This implies that their surfaces would be exposed to energetic particles from the star without shielding due to a strong magnetosphere as it is the case for the Earth. Massive planets near their host stars are much easier to detect, and several examples are known of such hot gas giants. The magnetic fields of Hot Jupiters, that means planets with masses $\geq \mathrm{M_J} = 1.89 \times 10^{27}$ kg, near their host stars could be only 1/10 in strength of that of Jupiter due to their slow rotation. It can be assumed, however, that due to strong stellar winds, these magnetospheres are compressed to a standoff distance at which the ionized part of the upper atmosphere builds an obstacle. The ionized atmosphere of such planets is on the other hand influenced by the XUV flux. Thus Hot Jupiters might have not been protected by their intrinsic magnetic fields (maybe when one neglects tidal locking because in the early phases of stellar evolution stellar winds had been much more intense). One can calculate neutral hydrogen loss rates of about 10^{10} g/s for the expanded exosphere of *HD 209458 b* (see [84].

1.1.4.2 Planets with Dense Atmospheres

A planet that is tidally locked to its host star is not exposed uniformly to the radiation of that star and therefore heated up unevenly. What happens when we consider such a tidally locked planet that has a dense atmosphere. Could it be possible that temperature inhomogeneities can be compensated? The present day Earth atmosphere contains about 350 ppm CO_2. Besides this carbon dioxide, water vapor is also a principal greenhouse gas. Let us assume a planet receiving Earth-level insolation but with an atmosphere containing 100 mbar CO_2 [100]. Then, this dense atmosphere would be able to transport enough heat flux to the dark side to prevent the atmosphere from freezing out there. It can even be shown that for a 15,000 mbar atmosphere and 0.8 Earth insolation, liquid water could be present all over the planet. In such an atmosphere, Heath [99] supposes that forest-habitable conditions can be expected.

There is one well-known example for tidal locking in extrasolar planetary systems: τ *Boo* is locked to the close orbiting giant planet τ Boo Ab. It was detected by high-precision spectroscopy measuring the motion of the host star. However,

also the Doppler shifted straylight reflected from the planet was detected (see [32]). *Gliese 581c* may be tidally locked to its parent star Gliese 581.

1.1.5 Tidal Heating

1.1.5.1 Io

It was a big surprise when on Io (Fig. 1.4), one of the Galilean satellites[4] of Jupiter, active volcanism was detected.[5] On Earth, the heat source that produces volcanic activity comes from its interior where radioactive materials decay and release energy and also from heat left over from its formation, accretional heat. This, however, cannot explain the volcanism on Io because the satellite is too small. The only explanation is by tidal heating. The semi-major axis of the orbit of Io is 421,800 km and one revolution around Jupiter takes only 1 d 18 h 27 min. The albedo of Io is 0.61, its diameter is 3643.2 km, and the mass is 8.94×10^{22} kg. The sidereal rotation period is exactly the same as is period of revolution.

On the surface of Io, 400 active volcanoes and more than 150 mountains have been detected. There are volcanic plumes that extend up to 500 km. The tidal forces of Jupiter on Io are 6,000 times stronger than those of the Moon on Earth. In addition

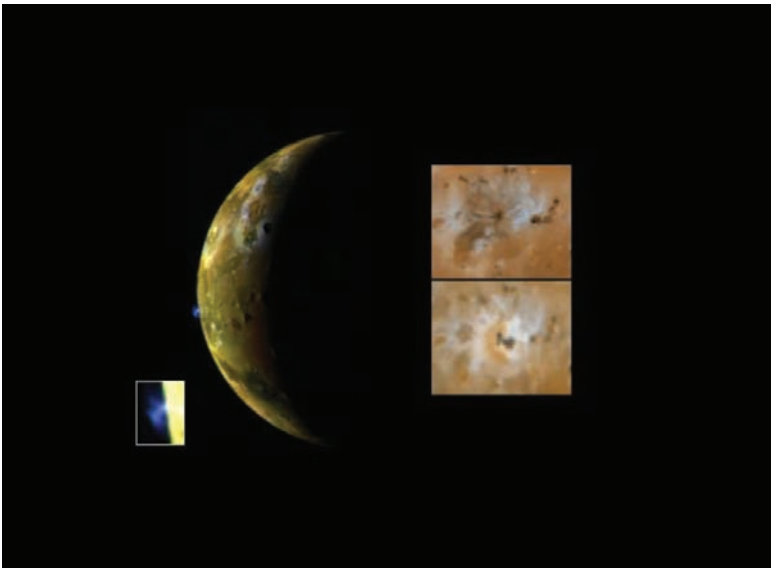

Fig. 1.4 The Jupiter satellite Io: due to tidal heating the body with the most active volcanism in the Solar System. Courtesy: NASA/Galileo mission, JPL

[4] They have been detected by Galilei, 1610.

[5] 1979, Voyager 1.

to these forces, there are also tidal forces caused by the two other Galilean satellites *Europa* and *Ganymede* (these are comparable to the tidal force of the Moon). Furthermore, the strength of these forces varies because the orbit of Io is elliptical. The variation of the tidal forces of Jupiter due to Io's elliptical orbit is 1,000 times the strength of the tidal forces of the Moon.

On Earth, the tidal force causes deformations of the whole Earth's crust between 20 and 30 cm. On Io, the deformations can reach up to 300 m.

1.1.5.2 Europa, Callisto

The effect of tidal heating due to Jupiter's enormous mass becomes even more interesting to our topic when considering two other Galilean satellites of Jupiter: Europa and Callisto. *Europa* (Fig. 1.5) has a mean orbital radius of 671,079 km and its orbital period is 3.55 days. The radius is 1,560.8 km and the rotation period is synchronous to the orbital period. The surface is extremely smooth and the albedo is 0.64 because the surface is composed of ice. On the surface, one observes dark streaks over the entire globe. It is supposed that they have been produced by a series of volcanic water eruptions or geysers. Since Europa is tidally locked to Jupiter, a certain pattern of these structures should have evolved. It seems, however, that this is only valid for the youngest patterns. The older the patterns, the greater the difference from this arrangement. This can be explained if Europa's surface rotates faster than its interior. A subsurface ocean could decouple the motion of the ice crust from that of the rocky mantle. The temperature at the surface of Europa is 110 K. But

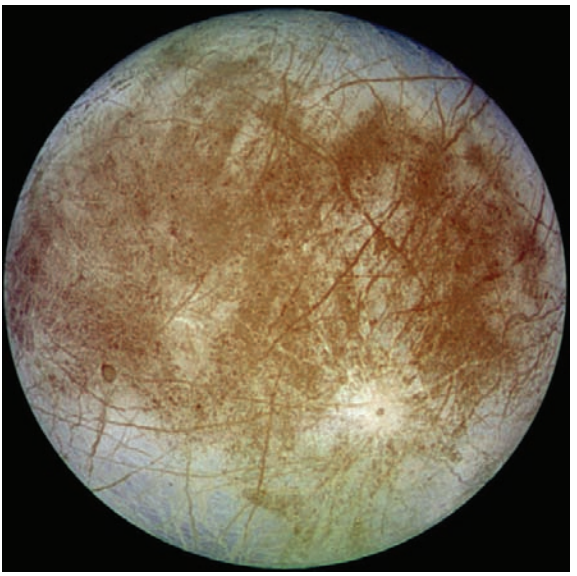

Fig. 1.5 The Jupiter satellite Europa: due to tidal heating beneath an ice curst an ocean of liquid water is highly probable. Courtesy: NASA/Galileo mission, JPL

due to tidal interaction, under the surface, there could be a liquid layer. The outer crust of solid ice is to be believed to be between 10 and 30 km thick. The underlying ocean could be about 100 km thick.

Jupiter's second largest satellite *Callisto* (Fig. 1.6) may also have a liquid ocean under its icy cratered crust. It is assumed that a salty ocean could exist there. The detection was made by measuring the changes of the magnetic field of Callisto (changes in Europa's magnetic field also gave a first hint to a liquid ocean). It was found that Callisto's field is also variable like the field of Europa. These changes can be explained by assuming varying electrical currents associated with Jupiter that flow near Callisto's surface. The only source where such currents can exist in Callisto is a layer of melted ice underneath. The result is that if this liquid layer were salty like Earth's oceans, then the electrical currents could be efficient enough to produce the field and the variations in it. One possibility would be that the water contains up to 5% of ammonia (see [233]). But there is one difference between Callisto and Europa. The mean radius of the orbit is 1.88×10^6 km. Thus, the tidal forces due to Jupiter are much smaller. Kuskov and Kronod [132] claim that Callisto must be partially differentiated into an outer ice, ice I-crust, a water ocean, a rock-ice mantle, and a rock-iron core mixture. They estimated the thickness of the ice I-crust between 135 and 150 km and that of the underlying water layer between 120 and 180 km.

Also, the possibility of a subsurface ocean on *Ganymede*, the largest satellite in the solar systems, is discussed.

Fig. 1.6 Callisto Courtesy: NASA/Galileo mission, JPL

1.1.5.3 Satellites of Saturn

Tidal heating is also important to explain the recently discovered thermal anomaly in the south polar region of the Saturn's satellite *Enceladus* (see Fig. 1.7). This moon is the sixth largest satellite of Saturn, its semi-major axis is 237,948 km, its orbital period 1.3 days, and its dimensions are $513 \times 503 \times 497$ km. The mean surface temperature is 75 K. Active eruptions have been detected and therefore sub-surface liquid water is supposed. Enceladus rotates synchronously with its orbital period and besides tidal heating an additional heat source results from its slightly elliptical shape and libration caused by resonance with other satellites of Saturn (e.g., Dione). However, it seems questionable whether an ocean could be stable there or not (see [209]). Sohl and Hussmann [231] discuss the probability of subsurface oceans on the Saturnian satellites and their conclusion is that for the largest satellite, *Titan*, an internal ocean is likely as well as for *Rhea*, whereas *Dione* and *Japetus* may have harbored oceans in the past because of the more intense radiogenic heat production. But it seems unlikely that *Tethys*, *Enceladus*, and *Mimas* had oceans at shallow depth.

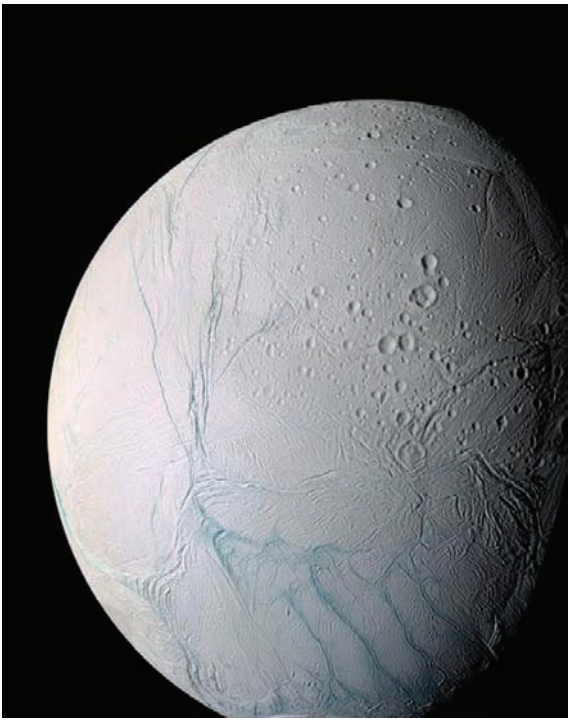

Fig. 1.7 Saturn's satellite Enceladus. The image was taken during the Cassini flyby 2005. Note that the surface features are similar to those on Europa. Courtesy: NASA

1.1.6 The Galactic Habitable Zone

After having discussed HZ in planetary systems (our own and extrasolar planetary systems), we investigate where the locations are in a galaxy that possibly contain stars with habitable planetary systems. Does the galactic environment have an influence on habitable zones?

→ The *galactic habitable zone* (GHZ) is the space in a galaxy where the conditions are favorable to life.

The universe started from a Big Bang about 14 billion years ago. All energy and matter was concentrated to a singular point and started to expand. At an age of about 3 minutes, the density and temperature of the universe was high enough so that thermonuclear fusion created about 25% He from the protons. This primordial fusion created the primordial composition of the universe: about 75% H and 25% He. All metals (elements heavier than He are called metals by astrophysicists) have been formed by nuclear fusion processes inside stars. Therefore, in the early cosmos there were no elements for the formation of planets like our Earth. The Sun is a star of second or even higher generation which means it formed from an interstellar cloud that was enriched by material already processed inside of stars and released when they exploded as a supernova.

In a galaxy, there is a gradient of the metal content: near the center, the metal content becomes higher than far out in the spiral arms. A planetary system must be located therefore close enough to the galactic center so that a sufficiently high level of metals exists. Only under these conditions, rocky planets could have formed. The necessity of heavier elements such as carbon in order to form complex molecules for life is also evident.

However, there are restrictions. If the planetary system is too close to the galactic center, then perturbations by passing stars become more likely. This will trigger hazards from comets that move into the inner planetary systems. Moreover, outbursts from supernovae and from the supermassive black hole at the center could cause strongly enhanced shortwavelength radiation bursts destroying the complex molecules needed for life.

Observations of extrasolar planetary systems have also shown that when the metallicity of the star becomes too high, more massive planets orbiting close to the stars are likely. These could have destroyed Earth-sized objects there.

The formation and expansion of the GHZ is illustrated in Fig. 1.8. In the early stages of galaxy formation, the heavy elements to form terrestrial planets were only present near the center of the galaxy because there the concentration of stars was largest and some stars already had evolved and became supernovae that enriched the interstellar medium with heavier elements there. This was not a safe environment because the danger of nearby supernovae explosions was high. Gradually, the heavy elements spread through the galaxy and terrestrial planets formed at greater and safer distances from the galactic center. The HZ appeared about 8 billion years ago at a certain distance from the galactic center. According to Lineweaver [145], the

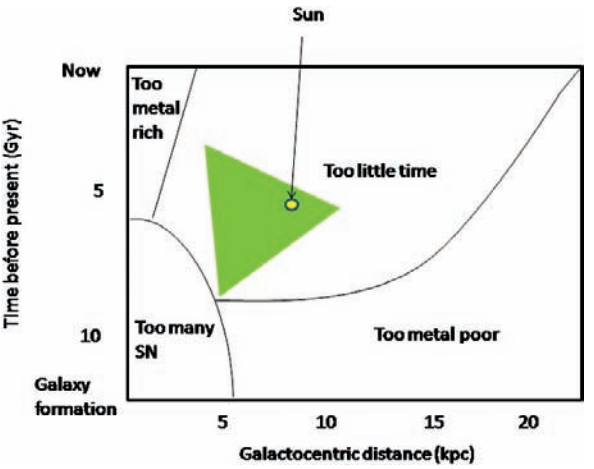

Fig. 1.8 The galactic habitable zone (GHZ) after [145]. The GHZ spreads out over the time. The Sun is located in the shaded area of the GHZ. Too close to the galactic center, SN explosions provide a hostile radiative environment; too far away, the metal content is not sufficient for terrestrial planets to form

GHZ is an annular region 7–9 kpc[6] from the galactic center. It widens with time and is composed of stars formed between 8 and 4 billion years ago.

Under these assumptions, an age distribution of stars can be derived and it is found that 75% of the stars in the GHZ are older than the Sun. If the evolution of life was on similar timescales as on Earth, then we belong to the "young generation" of galactic civilizations.

1.2 The Earth – Protector of Life

In this section, we give an overview about the conditions on Earth that enable life on our planet. We are protected against short wavelength radiation (mainly from the Sun) by the atmosphere and against particles (cosmic rays, part of it is also arriving from the Sun) by the Earth's magnetic field.

1.2.1 The Atmosphere

The chemical composition of the atmosphere is given in Table 1.1.

There is a high concentration of oxygen in our atmosphere. Under most planetary conditions, oxygen would react with other chemicals and the crust as was mentioned

[6] pc denotes parsec, 1 pc = 3.26 light years or 32.6×10^{12} km.

Table 1.1 Present day composition of the Earth's atmosphere

Element	Abundance
Molecular nitrogen	78%
Molecular oxygen	21%
Argon	1%
Water vapor	∼1% (highly variable)
Carbon dioxide	0.037% = 370 ppm
Methane	0.00015% = 1.5 ppm

above. The oxygen in our atmosphere is supplied continuously by photosynthesis. Also the 1.5 ppm methane component is not in chemical equilibrium with the atmosphere and crust and can only be explained by biological origin.

This composition of the atmosphere is nearly constant to an altitude of about 100 km. The highest variable compounds are water vapor and ozone. At higher altitudes, the composition of the atmosphere becomes inhomogeneous.

1.2.1.1 Troposphere

The troposphere extends up to 20 km at the equator and only up to 7 km at the poles. In the lower part of it, friction occurs with the Earth's surface. This part depends on the landform and is separated from the capping inversion layer. The troposphere is the densest part (75% of the total mass) of the atmosphere and here, nearly all weather processes occur and almost all water vapor and aerosols are found. The temperature drops from about 17°C to –50°C. Because air at the surface is compressed by the weight of all the air above it, the pressure of the atmosphere decreases with higher altitude. The change in pressure follows from the hydrostatic equilibrium:

$$\frac{\mathrm{d}p}{\mathrm{d}h} = -\rho g = -\frac{pg}{R_g T} \tag{1.16}$$

where ρ is the density, g the gravity, h the height, p the pressure, $R_g = 8.31\,\mathrm{J mol^{-1}\,K^{-1}}$ gas constant, and T the temperature. Assuming that g is constant, which is valid for low heights, we find that pressure decreases exponentially with height:

$$p(h) = p(h=0)e^{-\frac{gh}{R_g T}} \tag{1.17}$$

The temperature decreases with height at an average rate of 6.5°C per 1,000 m. This is due to *adiabatic cooling*: hot air rises and the atmospheric pressure falls, the air expands, temperature decreases (the air must do work on its surroundings).

In the troposphere, there is also a large *circulation system* (Fig. 1.9) that consists of three convection cells in each hemisphere: *Hadley cell*, *Ferrel cell*, and the *Polar cell*. These cells transport heat from the equator to the poles. The Hadley cell is a closed circulation loop: warm, moist air is lifted in equatorial low pressure areas to the tropopause (the border of the troposphere) and carried poleward. At about 30° N/S latitude, it descends in a cooler high pressure area. Some of that descending

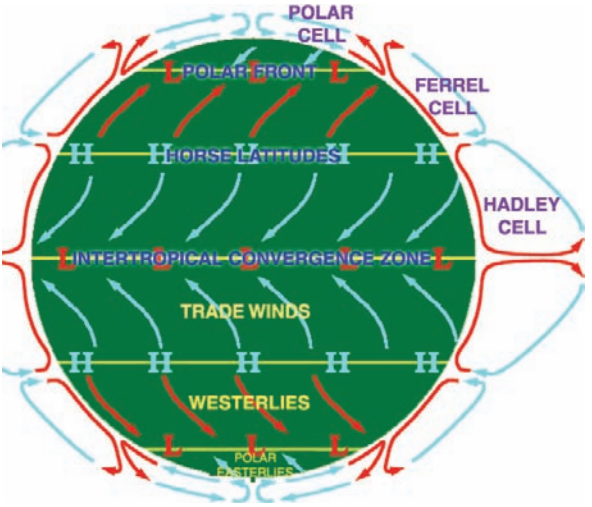

Fig. 1.9 The atmospheric circulation on Earth

air travels equatorially along the surface and creates the *Trade Winds*. The exact location of the Hadley cell depends on the Sun's zenith point that varies in the course of 1 year (thermal equator). The explanation of the Polar cell is likewise: at about 60° N/S latitude warm air rises and moves poleward to the N and S poles. During its travel to the polar areas, it cools and descends as a cold, dry high pressure area. The Ferrel cell lies between the Polar cell and the Hadley cell. Its explanation is more difficult and the reader is referred to specific textbooks.

These three cells are the major causes for global heat transport in the Earth's atmosphere. There are also other important factors like *latitudinal circulation*, which is caused by the different incident solar radiation (highest at the thermal equator). *Longitudinal circulation* arises because water has a higher specific heat capacity than land. The Pacific-Ocean cell is entirely ocean based. Normally, western Pacific waters are warm and the eastern Pacific waters are cool. This leads to a wind pattern, which pushes Pacific water westward and piles it up in the western Pacific (water levels in the western Pacific are about 60 cm higher than in the eastern Pacific).

1.2.1.2 The Stratosphere

It lies above the troposphere and extends to about 50 km. Motions in the stratosphere, are oriented horizontally. The temperature increases in the stratosphere. In the upper stratosphere, solar UV radiation splits the oxygen molecules $O_2 + h\nu \rightarrow O + O$ and the oxygen atoms react with the remaining O_2 to form ozone, O_3, creating the shielding *ozone layer*. The temperature at the top of the stratosphere reaches –3°C.

1.2.1.3 Mesosphere

The temperature decreases from the relatively warm stratopause to the mesopause (80–90 km altitude). The minimum temperature, –93°C, at the mesopause varies with seasons.

Stratosphere and mesosphere are called the middle atmosphere.

1.2.1.4 Thermosphere and Exosphere

Above 80–90 km altitude, temperature rises again as well as the number of charged particles because shorter than UV wavelength radiation gets absorbed by ionization processes. Therefore, this layer is also called the *ionosphere*. The extension of the thermosphere is up to 600 km and temperatures can reach up to 1,700°C.

At the top of the thermosphere starts the exosphere until it merges with interplanetary gases. Here, H and He are the prime components (extremely low density).

1.2.1.5 Atmosphere and Habitability on Earth

A relatively dense atmosphere is needed to provide low contrasts between day and night temperatures. The climate system on Earth is determined by complex interaction of wind systems, ocean currents, composition of the atmosphere, and solar insulation. The protective layers for life in the atmosphere are (i) the stratosphere, where UV radiation is absorbed and (ii) ionosphere, where EUV- and X-rays are absorbed by ionization processes.

1.2.2 The Magnetosphere

1.2.2.1 The Shape of the Magnetosphere

The physics of the plasma around astronomical objects such as planets, satellites, and stars is dominated by their magnetospheres; however, not all bodies possess a magnetosphere. To understand the physics and nature of the Earth's magnetosphere, we have to take into account its interaction with the interplanetary magnetic field and the solar wind particles. Its form is determined by the *solar wind*. On the side facing the Sun, it extends to 70,000 km (10–12 R_E, $1R_E = 6,371\,\text{km}$; all distances are given from the Earth's center). The boundary of the magnetosphere, *magnetopause*, is roughly bullet shaped. Its distance is $15\,R_E$. On the night side, there is a large extension in the magnetotail which looks like a cylinder with a radius of 20–25 R_E. The tail region stretches well past 200 R_E (Fig. 1.10).

The internal field of the Earth is generated in the liquid Earth's core by a *dynamo process* and the field is mainly a dipole field inclined about 10° to the rotation axis

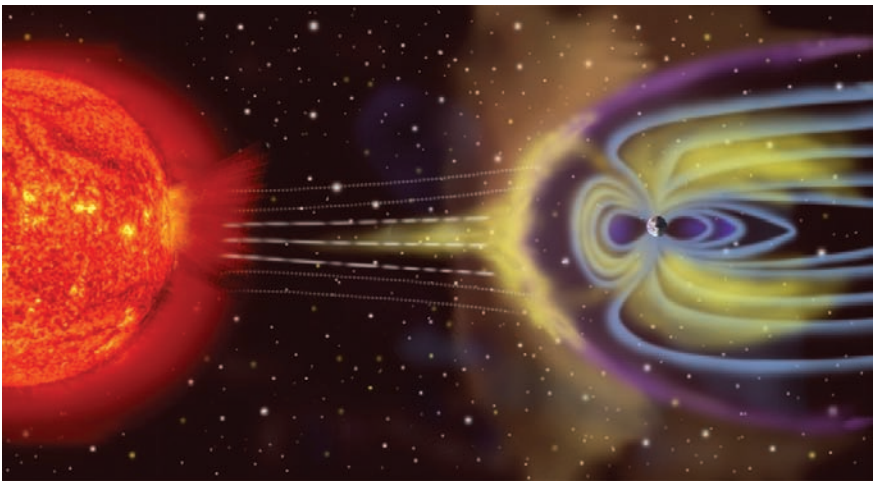

Fig. 1.10 The Earth's magnetic field protects us from the solar wind particles. Courtesy: NASA

of the Earth. The intensity of the dipole field is about $30-60\times10^3$ nT at the Earth's surface. Its intensity varies with the inverse cube of distance $\sim 1/R^3$.

1.2.2.2 Solar Wind and Magnetosphere Interaction

The solar wind is emitted from the solar corona (the velocity at the solar equator is about 400 km/s, at the Sun's poles up to 800 km/s) and consists of protons (95%), He-nuclei (4%), and heavier elements (C, N, O...) and, of course, electrons. At the distance of the Earth, the density is 6 ions/cm^3. It also contains a magnetic field (called *interplanetary magnetic field* (IMF)) of strength 2–5 nT. These are stretched out magnetic field lines from the Sun. The plasma from the Sun and from the Earth are separated by the magnetopause and, in principle, cannot penetrate. This is the shielding of our magnetosphere against the plasma flow from the Sun. But there occurs *magnetic reconnection* that means field lines of opposite sign (from the IMF and and the Earth's magnetic field) cancel and thus energy of the solar wind is transmitted to the magnetosphere causing strong disturbances. *Magnetic substorms* occur predominantly when the north–south component B_z of the IMF is large and points southward, which means antiparallel to the S–N direction of the Earth's magnetic field. Then reconnection is more likely with the Earth's magnetic field (Fig. 1.11). These processes are enhanced when solar activity level is high.

On the sunward side, at 13.5 R_E a collision-free bow shock occurs. The velocity of the solar wind particles is normally 2–3 times the velocity of the Alfvén waves[7] and in the region behind the shock wave (magnetosheath), the velocity drops to the Alfvén velocity.

[7] These waves are typical for disturbances that propagate in a magnetized fluid.

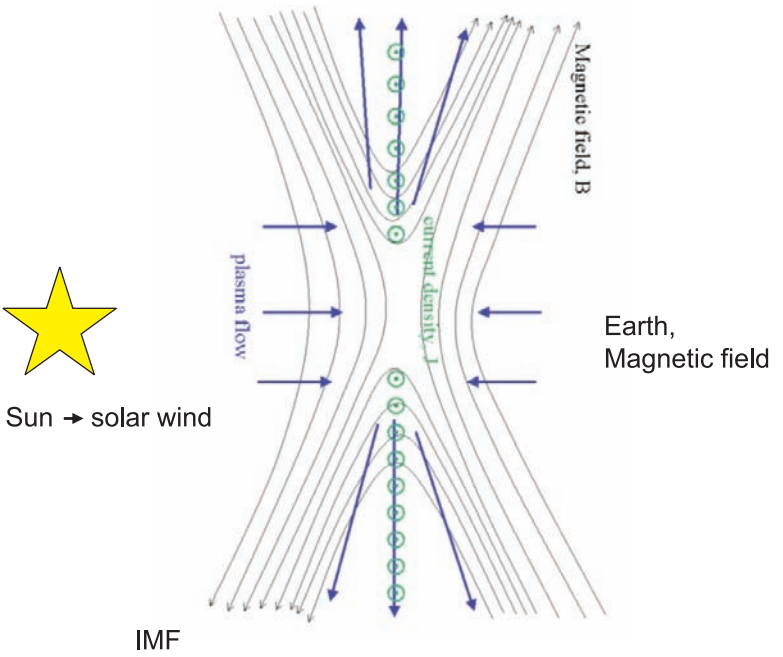

Fig. 1.11 Reconnection: the IMF is oriented antiparallel to the Earth's magnetic field

Satellites also detected the radiation belts:

- Inner belt: contains protons with energies 10–100 MeV, about $1.5\,R_E$
- Outer belt; at $2.5–8\,R_E$. The high-energy-part (>1 MeV) is called the outer radiation belt; the lower energy part (peak about 65 keV) is called ring current plasma

The outer belt and ring current are less persistent, the magnetic barrier can be broken by electric fields. These are generated by strong plasma perturbations.

1.2.3 Stability of Atmosphere and Magnetosphere

As we have seen, the Earth's atmosphere and the magnetosphere provide an effective shielding against radiation and charged particles. The question is how stable these systems are and have been over the past and what would be the consequences for habitability on Earth if this protection gets considerably weaker.

The topic was also recognized by Hollywood producers and a movie appeared in which the collapse of the Earth's magnetic field leads to massive electrical storms, blasts of solar radiation, and birds incapable of navigation. Indeed, it is well known that many animals on the Earth like birds, turtles, or bees rely on the geomagnetic field to navigate.

1.2.3.1 Magnetosphere — Reversals

The magnetosphere is known to reverse at irregular intervals ranging from 10^4 to 10^6 years. There seems to exist an average interval of approximately 250,000 years. The last reversal occurred 780,000 years ago and is named the *Brunhes-Matuyama reversal* (Fig. 1.12). This event is also useful in dating ocean sediment cores.

Past magnetic field orientation can be reconstructed since they appear frozen in the ferromagnetic minerals of solidified sedimentary deposits or cooled volcanic flows.

The reversal of the geomagnetic field over the past 160 million years is shown in Fig. 1.14. Note that there seems a quiet "Jurassic" zone as well as a Cretaceous Superchron. The Kiaman Superchron is another long time period during which there occurred no reversals of the geomagnetic field. The explanations why these reversals

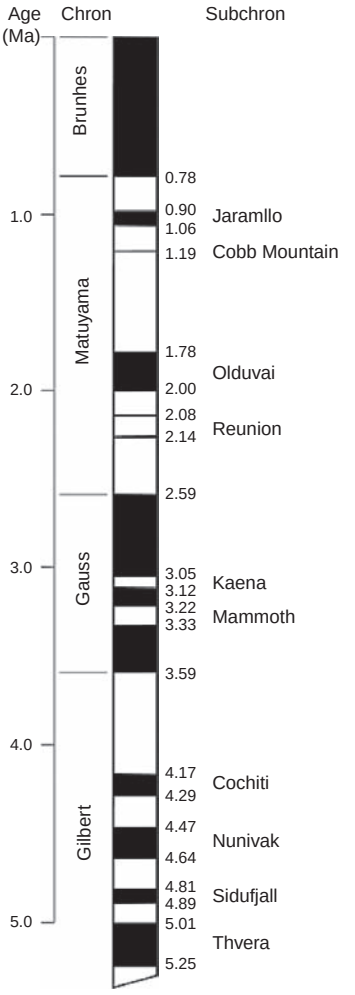

Fig. 1.12 Magnetic field reversals

occur are not definite. We know e.g. from the solar dynamo that these reversals occur (solar cycle, 22 years magnetic cycle). Reversals on Earth coincide with periods of low field strength. The reversals could be also triggered by the following:

- Plate tectonics
- Continental slabs that are carried down into the mantle
- Initiation of new mantle plumes
- Very large impact events

There are also brief disruptions which do not result in reversal. These are called *geomagnetic excursions*.

At present, the strength of the Earth's magnetic field is fading again. Today, it is about 10% weaker than it was in 1845 (at that time C. F. Gauß measured its strength). This will most probably not lead to a reversal during which the field more or less completely vanishes as it occurred 780,000 years ago. However, it is also known that over the southern Atlantic Ocean, a continued weakening of the field has diminished the shielding effect there. Satellites in low Earth orbit are vulnerable to radiation in this so-called *South Atlantic anomaly*.

The complex field structure during a magnetic field reversal is shown in Fig. 1.13.

One of the basic studies for geomagnetic field reversals was made by Parker [186]. Based on the fact that the magnetic field of Earth is steady with its present strength of about 0.5 Gauß[8] for periods of $10^5 - 10^7$ years and reverses abruptly in

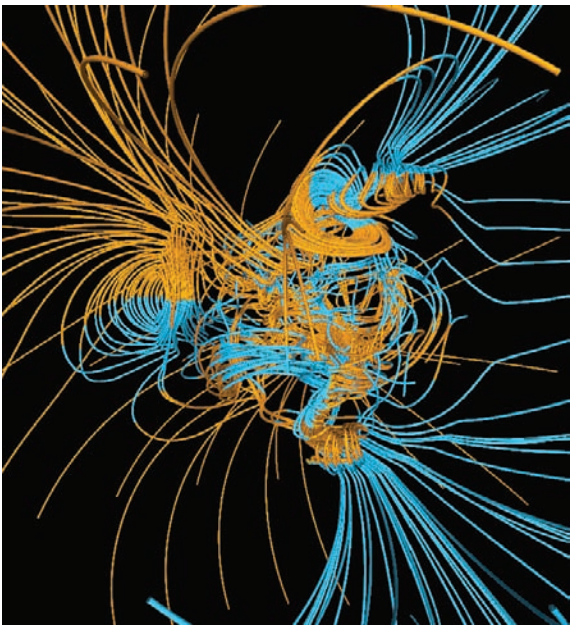

Fig. 1.13 Complex structure of magnetic field during field reversal. Adapted from G. A. Glatzmaier

[8] 1 Gauß $= 10^{-4}$ T, Tesla.

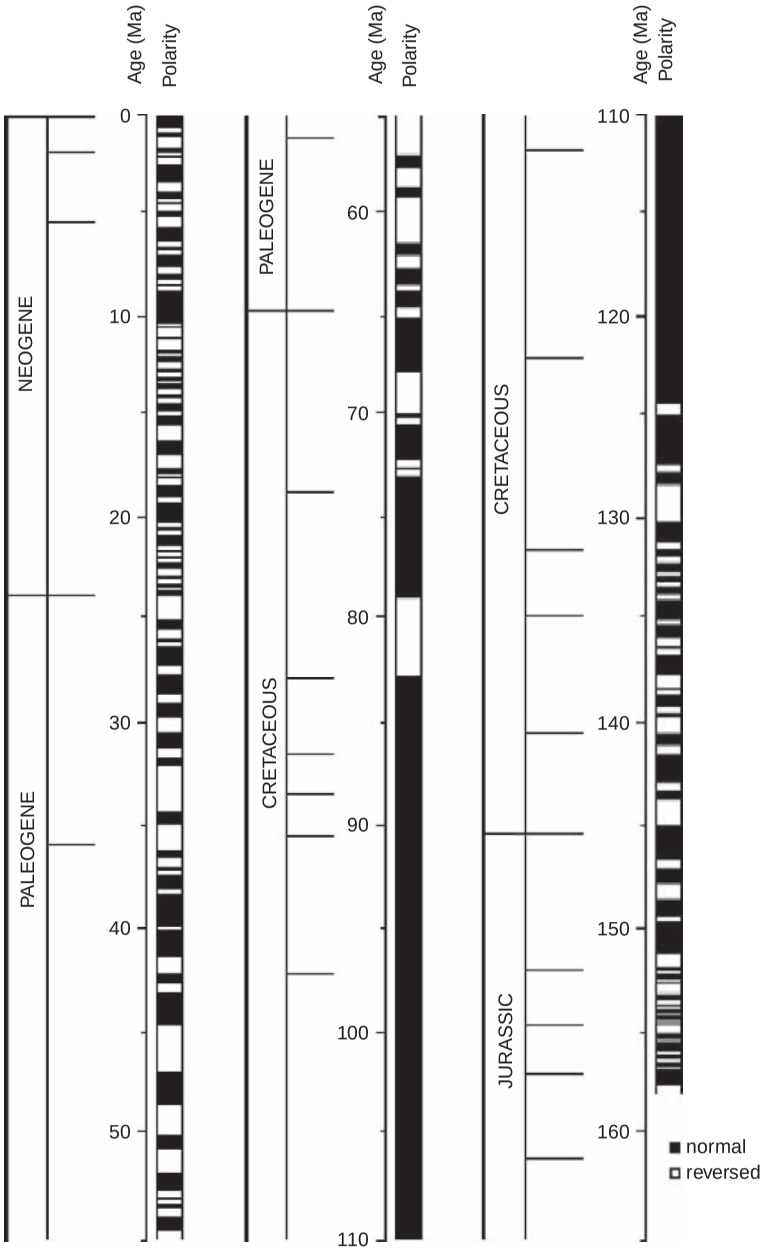

Fig. 1.14 Magnetic field reversals over the past 160 million years. Adapted from Lowrie, 1997, Introduction to Geophysics

times of 10^4 years, he concludes that a statistical fluctuation in the distribution of the 15–20 cyclonic convective cells in the core produces an abrupt reversal of the geomagnetic field. Thus, a reversal of the field can occur at random intervals.

Valet and Meynadier [250] used long-term records of geomagnetic field intensity and found two modes: one of field regeneration (timescale a few 10^3 years) and one of stable polarity stages (timescale $\sim 5 \times 10^5$ years), which is a slow relaxation process.

Glatzmaier and Roberts [74] used a three-dimensional self-consistent computer simulation. They were able to produce a stable geodynamo for over 40,000 years with several polarity excursions. At the end of the simulation, a reversal of the dipole was observed.

1.2.3.2 Mass Extinction During Field Reversals?

Let us briefly discuss the question here – what would be the consequences of such field reversals to the habitability of the Earth? Cosmic ray effects and faunal extinctions at geomagnetic field reversals were discussed by Black [21]. He found that a group of Radiolarian species became extinct simultaneously with the last field reversal. Because of the low magnetic field strength during the reversal time, cosmic rays would have been much more intense as well as the influences from solar flares and solar wind.

Raup [202] also studied whether there exists a correlation between magnetic field reversals and mass extinctions. His studies were based on the assumption that there is a periodicity pattern of 30 Myr for pulses of increased reversal activity. These intervals are centered on 10, 40, 70,... Myr BP. A correlation between reversal intensity and biological mass extinction was not found.

Doake [51] has studied whether there exists a statistical correlation between climatic change and geomagnetic field reversals. He used data for the Upper Pliocene period. He found that with a probability of 0.4 a climatic event will cause a magnetic field reversal.

Summarizing these findings, we may state that there have been reversals of the Earth's magnetic field in the past but there is no indication that these reversals lead to catastrophic changes on the Earth's habitability. The role of increased radiative and particle flux during times of low magnetic field strengths on the mutations, however, remains to be clarified.

1.3 The Habitable Zone in the Solar System

1.3.1 Mean Surface Temperatures

As we have shown in the previous sections, the HZ can be defined as the zone in which planets are found with surface temperature and pressure adequate so that

water can exist in liquid form. Let us do some simple case study: we assume planets at the positions 0.7–1.5 AU from central host stars with luminosity 0.7, 1.0, and $2.0L_{\odot}$. This luminosity range corresponds to the faint luminosity of the early Sun ($0.7\,\mathrm{L}_{\odot}$), the present day Sun value (1.0), and the luminosity of a red giant Sun nearly at the end of their evolution (about 4.5 billion years from now). In Fig. 1.15, the results for the faint young Sun are shown by the full line, the results for the present day Sun by the dotted line, and for the red giant phase Sun by the dashed line.

The figure shows that at present, the surface temperature on Earth at 1 AU is 255 K and it was 233 K at the time $L = 0.7\,\mathrm{L}_{\odot}$ and will raise to 302 K when $L = 2.0\,\mathrm{L}_{\odot}$. Of course, these values are not the correct surface temperatures because the effect of the Earth's atmosphere (natural greenhouse effect) is not taken into account. But it tells us that the temperature on early Earth would have been 22° lower than today. It is also clearly seen that the HZ progresses outward during stellar evolution. At the red giant phase of a solar-like star, an object located at 1.4 AU would have a similar surface temperature than now at 1 AU.

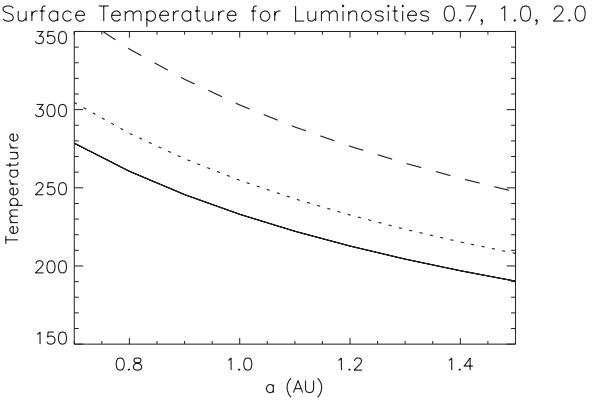

Fig. 1.15 Case study: surface temperature on a planet at different positions from the Sun versus solar irradiation at present day values (dotted line), at $L = 0.7\,\mathrm{L}_{\odot}$ (*full line*) and at $L = 2.0\,\mathrm{L}_{\odot}$ (*dashed line*)

1.3.2 Habitable Zone Around Giant Planets

At present, the solar system's HZ is just between the orbits of *Venus* and *Mars* with the Earth lying near the center of it. However, as it has been discussed above, this view of a HZ must be enhanced. Satellites around giant planets may well be in a HZ around them. There seems to be the following restrictions for a HZ *around giant planets*:

- The satellite must be large enough.
- The satellite must be heated by the giant planet – tidal heating (see the discussion of Io) provides the energy necessary for water to be in the liquid state.

- The giant planet's magnetic field provides a shielding against energetic particles.
- The giant planet must be at a certain distance to its host star, which is the case for the giant planets in the solar system.

Williams, Kasting, and Wade [269] studied this problem and we cite some of their results here. One problem is how satellites of giant planets obtained their volatiles. For the Earth, volatiles mostly have been provided from impacts of comets or carbonaceous asteroids. For satellites around giant planets far from their host star, such a scenario becomes difficult to explain. What happens when giant planets migrate closer to the host star? Will the satellite orbits remain stable? It has been proposed that *Titan*'s atmosphere was caused by such impacts, whereas the *Jovian Galileian satellites* have no atmospheres. Because of Jupiter's greater mass, the erosion of a possible atmosphere surrounding the Galilean satellites was much higher than for Saturn (also because of the sputtering caused by energetic particles from Jupiter's intense magnetosphere). Therefore, there might exist an upper limit for the giant planet's mass to host satellites that are habitable. A moon orbiting within the inner magnetosphere would lose its atmosphere due to such sputtering processeswithin $\sim 10^8$ years. The electron flux can be estimated at $4 \times 10^8\,\mathrm{cm^{-2}s^{-1}}$.[9] If the moon has a magnetic field, this effect will, however, be much less. It was measured that *Ganymede* has a magnetic field (its mass is only $0.03\,\mathrm{M_e}$).

Tidal locking is also important to consider. The *tidal locking radius* can be given as

$$r_T = 0.027 \left(\frac{P_0 t}{Q} \right)^{1/6} M^{1/3} \tag{1.18}$$

Here, M is the mass of the planet, P_0 the primordial spin period, t the time, and Q the so-called tidal dissipation factor. If we assume $P_0 \sim 15\,\mathrm{h}$, a time $t = 4.5 \times 10^9\,\mathrm{yr}$, and $Q \sim 100$ then the tidal locking radius for Jupiter becomes $r_T \sim 96\,\mathrm{R_j}$. This corresponds to a 116-day period. For a tidally locked satellite, the surface structure (ocean for example, shielded by ice) must compensate for such large temperature variations that result from these orbital and rotational periods.

On Earth, *plate tectonics* has played a major role for stabilizing the climate: subduction and metamorphism of carbonate sediments and return of CO_2 to the atmosphere. To drive plate tectonic, *radiogenic heat* must be available. The radiogenic heat flux can be estimated as

$$F_{\mathrm{rg}} \sim R \rho e^{-\lambda t} \tag{1.19}$$

where $\lambda = 1.5 \times 10^{-10}\,\mathrm{yr^{-1}}$ the decay constant for $^{238}\mathrm{U}$, R is the moon's radius, and ρ the moon's density. *Mars* became tectonically inactive 2 Gyr ago and the heat flow was about $40\,\mathrm{erg\,cm^{-2}\,s^{-1}}$. A moon that has the same density as Io would need a critical mass for tectonic activity over a long time of $\sim 0.23\,\mathrm{M_e}$.[10]

However, because of tidal heating, this constraint might be less important. Tidal heating occurs at a rate of

[9] For comparison: the solar-wind particle flux at the orbit of Mars is a factor of 10^3 less.

[10] $M_e = 5.97 \times 10^{24}\,\mathrm{kg}$, mass of Earth.

$$F_{\text{tid}} = (9/19)\rho^2 n^5 R^5 e^2 \frac{1}{\mu Q} \tag{1.20}$$

where n is the moon's mean motion about the planet, μ is the moon's rigidity (6.5×10^{11} dyn cm^{-2} for Io), and e the eccentricity of the moon. From this, the tidal surface heating of Io can be calculated as ~ 41 erg cm^{-2} s^{-1}.

All these considerations led to the conclusion that a $0.12\,M_e$ moon in an Io-like orbital resonance and possessing a Ganymede-like magnetic field could remain habitable for several billions of years. Therefore, systems belonging to *47 UMa* and *16 Cyg B* could have possible moons about their giant planets as candidates for extraterrestrial life.

According to these criteria, Reynolds, McKay, and Kasting [206] defined a zone around a giant planet in which larger satellites that are tidally heated exit a tidally heated habitable zone. In contrast to that, planets that are heated from a central star could be in a radiatively heated HZ.

Chapter 2
Properties and Environments of Life

In this chapter, we try to answer two basic questions: (i) how can life be characterized; what is the difference between nonliving and living substances? and (ii) how did life originate on Earth? Up to now, the Earth is the only planet where we can study life. We assume that life can only arise under similar circumstances that were found on early Earth. All objects (planets but maybe also large satellites of planets) that fulfill this criterion are called habitable.

The chemical background that is needed to understand basic mechanisms that are found in living cells is presented in the appendix, as well as a short introduction to the nomenclature of organic chemistry. Here, we discuss the occurrence of organic compounds in space, how life might have evolved from nonliving, and the structure of the cells and primitive living forms from the biological point of view. Finally, we briefly describe how primitive life forms on Earth have evolved. Bacteria, viruses, and cyanobacteria are discussed in more detail because these are the simplest form of life and they could probably be found elsewhere in the solar system as it will be discussed later. For the description of the biological processes, we followed [179, 136] and more details can be found there.

2.1 Organic Compounds in Space

Basic elements of organic chemistry are given in the appendix. In this section, we concentrate on organic compounds that can be found in space.

2.1.1 Interstellar Medium

The space between stars is not empty but it is filled with atoms and molecules and this mixture is called the *interstellar medium*. The physics of this mixture is quite different from that we know under laboratory conditions on Earth. In general, the interstellar medium is very dilute with densities from a few hundred up to about 10^8 particles cm^{-3}. The interstellar medium can be cool (temperatures below 100 K)

A. Hanslmeier, *Habitability and Cosmic Catastrophes*, Advances in Astrobiology and Biogeophysics, DOI 10.1007/978-3-540-76945-3_2,

or hot (temperatures up to 10^6 K). The interstellar medium consists of a gas component (mostly hydrogen) and about 1% of the interstellar medium is dust (micron sized particles, this can be inferred from polarization measurements). In total the interstellar medium contributes few percent to the total mass of a typical galaxy.

Interstellar molecules [27] can be observed at radio-, millimeter-, submillimeter- and IR-wavelengths. From these observations, many organic molecules were detected. It must be stressed that their abundance, relative to hydrogen (molecular), is less than 10^{-4} and in many cases even less. However, even very complex molecules like $HC_{11}N$ and diethyl ether, $CH_3-CH_2-O-CH_2-CH_3$, have been detected.

In the cold phases of the interstellar medium, there are icy dust grains.[1] These particles are important because they may act as catalysts.

The interstellar medium may also appear as dark clouds that absorb starlight. The densities are several 10^2 particles cm^{-3}. In such clouds, polycyclic aromatic hydrocarbons (PAHs), fullerenes, carbon-chains, diamonds, etc., are found.

Some of the molecules that have been detected in the interstellar medium are listed in Table 2.1.

Note that this table is only a sample of the list of molecules that have been detected. The detection of Glycine, H_2NH_2CCOOH, which is one of the 20 amino acids found in proteins was claimed by Kuan et al. [131] in the hot molecular cores Sgr B2(N-LMH), Orion KL, and W51 e1/e2 and later critically reviewed by Snyder et al. [230]. Cunningham et al. [49] used the Mopra Telescope[2] to search for glycine and the simple chiral molecule propylene oxide in the molecular clouds Sgr B2

Table 2.1 Some important molecules detected in the interstellar medium

CO	Carbon monoxide
FeO	Iron (I) oxide
H_2	Molecular hydrogen
HO	Hydroxyl radical
NO	Nitric oxide
NaCl	Sodium chloride
C_2H	Ethynyl radical
CO_2	Carbon dioxide
H_2C	Methylene
H_2O	Water
H_2S	Hydrogen sulfide
NH_2	Amino radical
SO_2	Sulfur dioxide
C_2H_2	Acetylene
H_2CO	Formaldehyde
NH_3	Ammonia
CH_4	Methane
HCOOH	Formic acid
CH_3OH	Methanol

[1] This has been investigated with the ISO satellite, Infrared Space Observatory.

[2] Mopra is a 22-m single-dish radio telescope located at the edge of the Warrumbungle Mountains near Coonabarabran, about 450 km north-west of Sydney.

(LMH) and Orion KL in the 3-mm band, but they were not successful and gave only some upper limits arguing that the detection of Kuan et al. was not real.

In 2004, a team detected traces of PAHs in an interstellar cloud, the most complex molecule, to that date, found in space. The use of PAHs has also been proposed as a precursor to the RNA world.

2.1.2 Organic Material Around Stars

The solar system is surrounded by a spherical shell that ranges from a distance of 50,000 AU to about 1 light year and contains billions of cometary like objects composed of ices like water, ammonia, and methane. The estimated mass of all objects in this Oort cloud is about 3×10^{25} kg. Such clouds are believed to occur also around other extrasolar planetary systems.

As we will discuss later, in comets and outer solar system bodies, organic molecules formed in the presolar nebula may have survived the solar system formation. Because of the low temperatures there, the molecules in the outer solar system have not been changed later. In the inner solar nebula, heating and thermochemical reactions were important and the molecules changed. Therefore, the outer solar system reflects the primordial state of the presolar nebula.

In comets and meteorites, many organic compounds are found. The Murchison meteorite fell in Australia on 28 September 1969. Seventy different amino acids were detected in this material. Initial analyses showed that the amino acids found there were racemic. A racemic mixture is one that contains equal amounts of left- and right-handed enantiomers (nonsuperimposable complete mirror images of each other) of a chiral molecule. A chiral substance is enantiopure or homochiral when only one of the two possible enantiomers is present. Biological homochirality evolved on a presumably racemic primordial early earth. Homochirality can, therefore, be regarded as a sign for biologic activity.

We will discuss the formation of planets from so-called protoplanetary disks in more detail later. Let us mention here that planetary systems form out from the gravitational collapse of a protoplanetary disk. Many organic compounds were found in such disks. Water is an essential component for life. Today, it is assumed that impacts provided a substantial fraction of organic compounds and even water, which is necessary for life to form on Earth.

2.2 Life

Some basics of organic chemistry are outlined in the appendix. Now, we want to discuss the question, what is life and offer some ideas of how life could have originated.

2.2.1 What Is Life?

This question is not a trivial one. Growth is not a criterion, crystals can also grow. Movement alone is not a valid criterion, e.g., when a pellet of dry ice (frozen carbon dioxide) is dropped into water, the pellet will move randomly across the surface.

Living organisms consist of *cells*, which are mainly composed of *cytoplasm* bound by a very thin *membrane*. All living cells contain genetic material. This controls their development and activities. In many cells, the genetic material, DNA, is found in the nucleus, a spherical structure suspended in the cytoplasm. In simpler forms of life such as bacteria, the DNA is distributed in the cytoplasm.
Besides consisting of cells, life has other "properties":

Growth: new cells are produced.
Reproduction: when organisms reproduce, the offspring resemble the parents.
Response to stimuli: a major characteristic of all living things. Plant responses to stimuli are generally much slower than those of animals.
Metabolism: collective product of all the biochemical reactions taking place within an organism. New cytoplasm is produced, damage repaired, and normal cells are maintained. Metabolism includes photosynthesis, respiration, digestion, and assimilation.
Movement: plants can also move. Leaves of sensitive plants (e.g., *Mimosa pudica*) fold within a few seconds after being disturbed or subjected to sudden environmental changes.
Complexity of organization: cells are composed of large numbers of molecules (typically more than 1 trillion in a typical cell). The molecules are organized into compartments, membranes, and other structures in the cell. Bacteria are considered to have the simplest cells known, yet this cell contains at least 600 different kinds of protein and 100 other substances. Other organisms are more complex.
Adaption to the environment: living organisms respond to their environment – to air, light , water, soil, etc. Natural selection leads to adaption to their environment. Today, many species are threatened with extinction because they are not able to adapt fast enough to the changing environment.

2.2.2 From Nonliving to Living

The key questions for the emergence of life can be summarized as follows:

- What are the basic building blocks and how were they synthesized first?
- Does there exist a type of primordial synthetic chemistry that leads to the above process?
- Where did self-assembly occur (e.g., areas of geothermal activity)?

2.2.2.1 Life on Early Earth

In 1903, S. Arrhenius [5] (1859–1927) suggested an extraterrestrial origin of life on Earth. He argued that life might have reached the Earth in the form of microscopic

spores that have been transported through space by the radiation pressure of stars. This is also called the *panspermia hypothesis*. At that time, Lowell discussed the Martian channels that were described first, when Mars was near to Earth (Opposition) in 1877, by Schiaparelli (he claimed to have seen linear structures on Mars which he named by the Italian word *canali*). Lowell and Flammarion suggested that these *canali* were constructed by intelligent Martians to distribute water on the dry planet but Arrhenius did not believe that. However, Arrhenius produced a map in which huge carboniferous swamps were assumed on Venus.

In 1920, A. Oparin and J.B.S. Haldane proposed that life on Earth might have originated from a reducing, primitive atmosphere composed of methane, ammonia, and water. The source of carbon for life was believed to be methane (Oparin) or carbon dioxide (Haldane). Both supposed that life first began in the atmosphere and then in the oceans. In 1982, A.G. Cairns-Smith proposed that mineral life-forms evolved from complex prebiotic clay crystals. Thus, there is no need for a reducing atmosphere or a primordial soup [31].

In the 1950s, it was assumed that the primitive Earth atmosphere consisted of methane (CH_4), ammonia (NH_3), hydrogen (H_2), and water (H_2O) and S.L. Miller and H.C. Urey [166, 167] carried out a famous experiment at the University of Chicago. They modeled the primitive Earth atmosphere and ran continuous electric currents simulating lightning storms, which were very common on the early Earth to this environment. After 1 week, 10–15 amino acids were found in the primordial soup.

Another experiment was conducted later in 1961 by J. Oro and he was able to produce amino acids from hydrogen cyanide (HCN) and ammonia in an aqueous solution.

However, today, there are strong objections against these experiments. The primitive Earth atmosphere contained much less reducing molecules. It did not consist mainly of NH_3 and CH_4 but most probably of $CO_2, CO,$ and N_2. There are two reasons to support this idea: first, volcanic eruptions contain these gases and second, solar UV radiation destroys NH_3 and CH_4. These molecules would have been short-lived. Solar UV photolyzes H_2O:

$$h\nu(\mathrm{UV}) + \mathrm{H_2O} \rightarrow \mathrm{H} + \mathrm{OH} \tag{2.1}$$

The OH radicals attack methane, which results in CO_2 and release H_2. The hydrogen will be lost into space.

It was found that gas mixtures of $CO, CO_2,$ and N_2 in the absence of O_2 give similar results as those used by Miller and Urey. It was possible to produce most of the natural amino acids, purines[3], and sugars using such mixtures. When oxygen was added, however, no organic molecules formed.

Sagan and Chyba [215] proposed that the early Earth had an organic haze layer in its atmosphere. Such a layer can be found in the atmosphere of Titan, the largest

[3] Purines are heterocyclic aromatic organic compounds. They consist of a pyrimidine ring fused to an imidazole ring. Purines and Pyrimidines make up the two groups of nitrogenous bases which are a crucial part of both deoxyribonucleotides and ribonucleotides, and the basis for the universal genetic code.

satellite of Saturn and is produced by methane photolysis in the presence of nitrogen. An organic haze layer would preferentially absorb ultraviolet light, thereby allowing ammonia and methane to persist in the atmosphere. However, as in the case of Titan, such a layer would also have an antigreenhouse effect [162], which could oppose or even cancel any greenhouse effect generated by the shielded methane and ammonia. The question of whether a combination of an organic haze layer and methane and ammonia gases can combine to produce a net greenhouse effect on the early Earth depends on the optical properties of the organic haze layer (cited from Imanaka, Khare and McKay [114]).

2.2.2.2 Hydrothermal Vents

Hydrothermal vents are openings, in the ocean floor. From these openings, hot, mineral-rich water escapes or erupts. In the 1970s, black smokers (Fig. 2.1) were detected. These are chimney-like structure above hydrothermal vents.[4] In these smoker chimneys, sulfides of iron, copper, and zinc are found. At the mixture of the hot mineral-rich water with cold water, these sulfides are precipitated and the vent water, therefore, appears black in color.[5] The most striking discovery was that these warm chemical rich environments are the living space for many species. Strange animals, like a 1-m-long reddish worm (*vestimentiferan*), which builds and lives in a tube up to 7 m long, were found. The chemoautotrophic archea use chemical energy derived from the breakdown of hydrogen sulfide to build organic compounds. Chemautotrophs are microorganisms (many archaea and bacteria) that synthesize organic compounds they need from inorganic raw materials in the absence of sunlight (i.e., no photosynthesis).

In the 1980s, G. *Wächtershäuser* [255] proposed his iron-sulfur theory. In contrast to the classical Miller experiments, which depend on external sources of energy (such as stimulated lightning or UV irradiation), the energy is released from redox reaction of sulfides of iron and other minerals. This energy permits the synthesis of organic molecules and also the formation of oligomers[6] and polymers. Huber and Wächtershäuser [111] modelled volcanic or hydrothermal settings. They showed that amino acids were converted into their peptides by use of precipitated (Ni,Fe)S and CO in conjunction with H_2S (or CH_3SH) as a catalyst and condensation agent at 100°C and pH 7–10 under anaerobic, aqueous conditions. Thus, a thermophilic origin of life seems plausible.

2.2.2.3 Extremophiles

Organisms that live under extreme physical or geochemical conditions are classified as extremophiles. Most of them are microbes. Some types of extremophiles are listed in Table 2.2.

---

[4] Sometimes also called underwater geyser.

[5] Hydrogen sulfide H_2S is the closest chemical analogue of water.

[6] Conversion of monomers to a finite degree of polymerization.

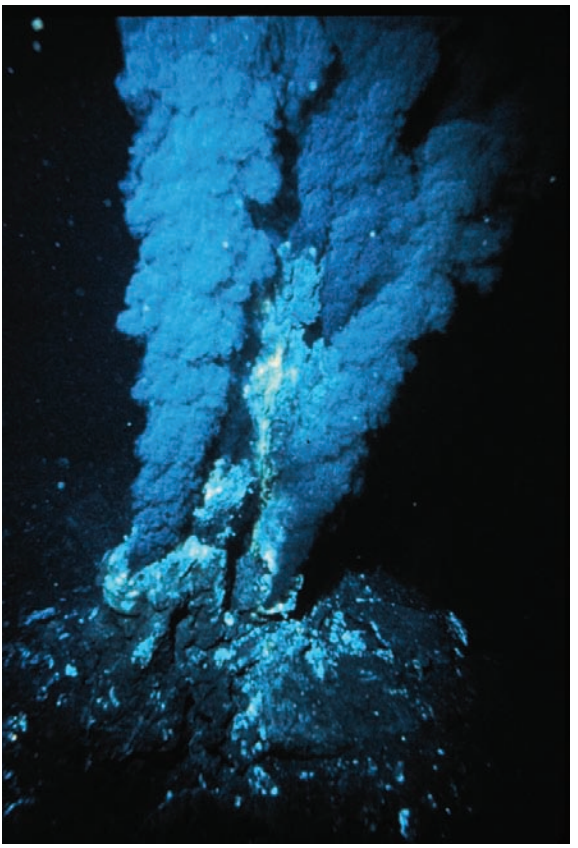

Fig. 2.1 Black smoker at a mid-ocean ridge hydrothermal vent. Credit: OAR/National Undersea Research Program (NURP); NOAA

The hyperthermophiles were discovered in the 1960s in hot springs in Yellowstone National Park. The hyperthermophile *Strain 121* is able to survive a temperature of 121°C and even to double its population within 24 hours in such an environment. It consists of hyperthermostable proteins. Another example is *Pyrolobus fumarii*, an Archaea living at 113°C in Atlantic hydrothermal vents.

The new findings about extremophiles are of particular interest for our study. If an asteroid impact on Earth caused the global temperature to decrease or rise considerably and most of the highly advanced species died in a mass extinction, then it is plausible that at least such extremophiles survived a cosmic catastrophe.

Extremophiles and extreme environments in the context of astrobiology and the search for extraterrestrial life were investigated by Cavicchioli [35].

2.2.2.4 Molecular Self-Assembly

There exist many examples of self-organization in biology: spontaneous folding of proteins and other biomacromolecules, formation of lipid bilayer membranes

Table 2.2 Some types of extremophiles

Name	Environment
Acidophiles	Live in pH ≤ 3
Alkaliphiles	Live in pH ≥ 9
Barophiles	Live under high pressure
Piezophiles	-"-
Endoliths	Live in microscopic spaces within rocks
Cryptoendoliths	Live in fissures, faults in deep subsurface
Halophiles	Need salt to grow
Hyperthermophiles	Live at T = $80-120^{\circ}$, hydrothermal vents
Hypoliths	Live inside rocks in cold deserts
Litoautotrophs	Source of carbon is CO_2
Metatolerants	Capable of tolerating high levels of metals like copper, cadmium, zinc.
Oligotrophs	Grow in nutritionally limited environment
Osmophiles	Grow in sugar concentration
Psychrophiles, cryophiles	Grow at T $< -15^{\circ}$; permafrost,polarice..
Radioresistants	Tolerate high levels of ionizing radiation
Thermophiles	Live at temperature up to 80°C
Xerophiles	Grow in extremely dry conditions

and also social animals like social insects (bees, ants, and termites), formation of flocks by birds, etc. Molecular self-assembly is crucial to the function of cells: self-assembly of lipids to form the membrane, the formation of double helical DNA through hydrogen bonding of the individual strands, etc.

Moreover, the transition to life from nonliving matter is strongly connected with molecular self-assembly. Nature has always worked bottom up, from the simplest assemblies to more complex structures. Molecular self-assembly can be defined as *spontaneous emergence of highly organized functional supramolecular architectures from single components of a system under certain external condition*.[7] To study the origin of life, molecular self-assembly[8] seems a crucial starting point. Hazen and Deamer [98] studied reactions of pyruvic acid under hydrothermal conditions. A complex mixture of larger organic molecules resulted, some of which are amphiphiles[9] that readily self-assemble into cell-sized vesicular structures.

An unsolved question is the obvious homochirality, which is a symmetry break in nature. A molecule is chiral if it cannot be superimposed on its mirror image. Homochirality means that molecules possess the same sense of chirality. All of the 20 biologically active amino acids are left handed. Amino acids that are produced in the laboratory are 50% left handed and 50% right handed. It is an interesting fact that life on Earth is based on left handed (L) amino acids and right handed (D) sugars.[10] It is even assumed that homochirality is a characteristic signature of life.

[7] After [140].

[8] See as a review: [274].

[9] A chemical compound possessing both hydrophilic and hydrophobic properties.

[10] The explanation maybe parity-violating energy difference.

Finding molecules of one chirality on a planet could be used as a signature for life or at least prebiotic chemistry (see also [47]).

In recent years, this fundamental problem for the origin of life was tested experimentally by direct microscopic verification of self-assembled DNA bases on mineral template surfaces. A sizzling technique was applied to simulate possible spontaneous self-assembly of two-dimensional molecular crystalline layers onto mineral surfaces with a primordial soup scenario. The free parameters of such experiments are the concentration of the solution, the substrate temperature (was tested up to 130°C), the quantity of the solution, and temperature+time tempering. The scenario gives a model for the assembly of polypeptides based on a 2d-DNA-base at the solid–liquid interface: inorganic template → organic template → chiral symmetry break → enzyme, protein, etc.[11]

It must be stressed that simple organic molecules are of course a long way from a fully functional self-replicating life form. The spontaneous formation of complex polymers from abiotically generated monomers is not straightforward. According to the "soup theory," these simple organic molecules played an important role and were accumulated by various processes.

A source for the complex molecules might have been impacts that were quite numerous on early Earth.

M. Eigen (1970s) [54] tried to examine the transient stages between the molecular chaos in a prebiotic soup and the transient stages of a self-replicating hypercycle, between the molecular chaos in a prebiotic soup and simple macromolecular self–reproducing systems. The hypercycle consists of a cyclic sequence of individual self reproducing cycles. These cycles consist of RNA and peptide molecules and there are interactions and back reactions between them. In this context, the term *self-organization* again becomes important.

Polymers in the range of 30–60 monomers are needed for a genetic system. The problem is how these can be formed because in aqueous solution hydrolysis competes with polymerization. Mineral surfaces to facilitate prebiotic polymerization [59].

2.2.3 Cells

Cells are the structural and functional unit of all known living organisms; some organisms ar unicellular, they consist of a single cell (most bacteria); humans consist of about 10^{14} cells.

Cells consist of cytoplasm and the structures within.

About 96% of the cytoplasm are composed of C, H, O, and N. Three percent consist of P, K, and S. The remaining 1% includes Ca, Fe, Mg, Na, Cl, Cu, Mn, Co, Zn, and minute quantities of other elements.

[11] Review given by Ref. [101]

2.2.3.1 Prokaryotic, Eukaryotic Cells

There are two basic types of cells:

- Cells without nuclei and some other features: *prokaryotic* cells.[12]
- *Eukaryotic* cells[13] possess a nucleus.

Most plant cells and animal cells are tiny and they are invisible to the unaided eye. Prokaryotic and eukaryotic cells have in common the cell membranes, cytoplasm, genetic material, and energy currency enzymes and coenzymes. All these are necessary to maintain a cell alive.

In biology, all organisms are classified into large categories called *domains*.

Most single-celled organisms are *prokaryotic bacteria*. There is another class called *archaebacteria*. Bacteria are a few micrometers long have many different shapes such as spheres, rods, or even spirals and they are found on quite different habitats on Earth: in soil, acidic hot springs, or even in radioactive waste and they can survive outer space (that means extreme cold and vacuum). In a gram of soil, there are typically 40×10^6 bacteria and in a milliliter of water, about 10^6. The ancestors of modern bacteria were single celled organisms that developed as first forms of life about 4×10^9 years ago. *Stromatolites* are fossils. They have been formed by sedimentation and cementation of microorganisms such as the cyanobacteria.[14]

The *archaea* are more closely related to eukaryotes than the bacteria. Their genetic transcription and translation is similar to those of eukaryotes. Most bacteria and eukaryotes have membranes that are composed of glycerol-ester lipids. The membrane of the archaea is composed of glycerol-ether lipids. Archaea are found in extremely hot environments (hyperthermophiles) but they can be also found in all habitats.

2.2.3.2 Cell Membranes

Membranes confine and protect a cell from the environment, and also the organelles inside the cell have membranes. The cell membranes are thin sheets and are composed of phospholipids and proteins. The phospholipid molecules are polar molecules- one end (the glycerol portion) is soluble in water (hydrophilic), the other end is not water soluble (hydrophobic) and is comprised of fatty acids. Phospholipid molecules form a double-layered sheet with the hydrophilic portions of the molecules facing away from each other.

Thus, in water, the lipophilic tails line up against one another forming a membrane with hydrophilic heads on both sides facing water. *Liposomes* are formed spontaneously. These are small lipid vesicles used for transport of materials into the cells and allow for diffusion.

---

[12] pro means before, karyon means nucleus.

[13] eu means well or good.

[14] An old term no longer used now is blue green algae.

The protein component of a cellular membrane can be found located either on the surface of the membrane or inside it. They help to transport molecules and also allow organisms to recognize differences between types of cells – disease causing organisms are recognized because their surface proteins are different from those of the organism's own cellular membranes. Cell membranes have molecules on their surface known as cluster of differentiation or CD markers and they are designated as CD1, CD2, ...CD4 cells and are found inside the human body (blood cells). [15]

In the membrane, cholesterol is also found, which helps in stabilizing the membrane and keeps it flexible.

2.2.3.3 Diffusion and Osmosis

One of the main activities of cells is the exchange of nutrients and wastes with the environment by diffusion, osmosis, facilitated diffusion, active transport, and phagocytosis.

The process of *diffusion* is well known from molecular kinetics. The motion of molecules is completely random. But when considering two types of molecules there is a trend to form a mixture (e.g., sugar in water). That is called the net movement and it depends on the relative concentration of the molecules, the concentration gradient (diffusion gradient). When the molecules are equally distributed, there is no net gradient. The permeability of molecules through the plasma membrane depends on the size and ionic charge as well as on the solubility of the molecules. The cell has little control and, therefore, diffusion is considered a passive process. The direction of diffusion is always from high to low concentration.

Osmosis denotes a selectively permeable membrane and the diffusion associated with it. The membrane allows, for example, the diffusion of molecules that are small (e.g., water) and are able to dissolve in phospholipids (e.g., vitamins A and D), but the passage of larger molecules is prevented. Osmosis is essential for the plants. If the water concentration outside the plant cells is higher than that inside, more water molecules enter the cell, which creates an internal pressure. Osmosis, like diffusion, is a passive mechanism because the cell does not control it.

One example of *facilitated diffusion* is the movement of glucose molecules throughout the cell. This is accomplished by specific proteins – active carrier proteins.

For a cell to function, a certain amount of water is necessary, but too much water may interfere with the chemical reactions, water is necessary to dissolve poisonous products in the cell.

Endocytosis (or phagocytosis) occurs when the cell wraps around a particle and engulfs it.[16] The material is brought into the cell in a vacuole and inside the cell, the vacuole breaks up; the contents are released and, e.g., destructive enzymes start to work.

[15] The CD4 acts as the attachment site for human immunodeficiency virus (HIV). The virus enters the CD4 cell and commands its metabolism but HIV cannot attach to cells that do not have the CD4 surface markers (e.g. skin cells).

[16] For example Leukocytes, white blood cells, they surround invading bacteria, viruses.

2.2.3.4 Organelles and Membranes

The *endoplasmatic reticulum* (ER) consists of folded membranes and tubes throughout the cell. These membranes form a large surface enabling chemical reactions to take place there.

In the *Golgi apparatus*, certain molecules are synthesized and chemical reactions are concentrated. The active or inactive forms of enzymes are controlled; lysosomes can decompose dead or dying cells, selectively destroy cells, digest macromolecules, and kill dangerous microorganisms.

The *nuclear membrane* separates the nucleus (nucleoplasm) from the cytoplasm. The nuclear membrane is formed like a sphere around the DNA.

The *mitochondrion* resembles a small bag. On the inner surface, there are particular proteins and enzymes, which enable aerobic cellular respiration. Aerobic cell respiration means that energy is produced from food molecules using oxygen. A typical human cell contains up to 10^4 mitochondria. Muscle cells contain even more. When there is activity in the cells, the mitochondria swell; when there is low activity, they shrink.

Chloroplasts are found only in plants and algae. There are cells that contain only one large chloroplast; others contain hundreds of smaller chloroplasts. In the chloroplasts, by *photosynthesis* light energy is converted to chemical-bond energy.

Other nonmembranous organelles in the cytoplasm include the *ribosomes*. They are composed of RNA and protein. The ribosomes manufacture protein. Liver cells, e.g., actively produce protein and they contain great numbers of free and attached ribosomes. Other nonmembranous organelles are the microtubules, microfilaments, and intermediate filaments. These provide the cell shape and support.

Chromatin is loosely organized DNA/protein strands in the nucleus. When the chromatin is tightly coiled into shorter, denser structures, we call them *chromosomes*. Besides chromosomes, the nucleus contains one to several nucleoli. The liquid matrix in the nucleus is called the nucleoplasm.

Finally, we mention that many cells have cilia or flagella. These are hairlike structures from their surfaces and the cell is able to control the action of these microtubular structures.

2.2.3.5 Cell Types

The prokaryotic cells (bacteria and archaea) do not have a typical nucleus bound by a nuclear membrane and they do not contain mitochondria, chloroplasts, Golgi, or extensive ER. On the other hand, prokaryotic cells contain DNA and enzymes. They are able to reproduce and show metabolism and perform all the basic functions of life. The archaea make up to 20% of the biomass. The first arachea were detected in hostile environments but later it became clear that they occur also in soil and in the oceans. There are no known members of archaea that cause diseases.

Only 5% of prokaryotic bacteria cause diseases (tuberculosis, gonorrhea, etc.). Bacteria are also involved in the decay and decomposition of dead organisms. Many

bacteria are surrounded by a capsule enabling them to stick to surfaces and resist phagocytosis. Most bacteria are rods (bacilli), spherical (cocci), or curved (spirilla).

The eukaryotic cells contain a nucleus and most of the other organelles described above. Cells of plants, fungi, protozoa, and animals are prokaryotic; the cells of plants and algae contain chlorophyll, thus they contain chloroplasts in their cytoplasm.

Fungi (yeasts, molds, and mushrooms) cells lack chlorophyll. They are not able to produce their own food (like the plants) and they have no cell walls like the animals. The protozoans (Amoeba, Paramecium) have all the organelles described above and must consume food as do the fungi and the multicellular animals.

2.2.3.6 Cell Division

By cell Division, dead cells are replaced, damaged tissues are repaired, and the growth of the organism is enabled. The human body loses about 50 million cells per second that must be replaced. In the process called *mitosis* (Fig. 2.2), the genetic information is equally distributed to two daughter nuclei. The stages of mitosis are the interphase, the prophase (replicated chromatin strands become recognizable as chromosomes, nuclear membrane fragments, and centrioles move to the cell poles), metaphase (chromosomes move to the equator of the cell – in a human cell there are 46 chromosomes aligned at the cell's equatorial plane), anaphase (DNA replication completed, daughter chromosomes contain identical information), and telophase (division begins, cytokinesis). Cancer refers to any abnormal growth of cells that has a malignant potential.

Each new cell gets one of the two daughter nuclei, both have a complete set of genetic information.

There are five stages in the life of a eukaryotic cell:

1. G_1 gap (growth),
2. S synthesis,
3. G_2 gap, growth – phase two,
4. Cell division (mitosis, cytokinesis),
5. G_0 gap (growth), dormancy or differentiation.

The first three phases – $G_1, S,$ and G_2 are also known as interphase. The cell is not dividing but engaged in metabolic activities (e.g., muscle cell contractions and photosynthesis). The individual chromosomes are not visible. The DNA strands of the chromosomes contain different information on enzymes required for digestion, how to produce enzymes that will enable metabolism and eliminate harmful wastes, how to repair cell parts, production of a healthy offspring, how to react to changes of the environment, and coordination and regulation of essential life functions.

Another important form of cell division is *meiosis*. Here, there are, in principle, two cell divisions. This results in cells with haploid number of chromosomes (haploid means one set of chromosomes). Germ cells (eggs and sperms) are produced

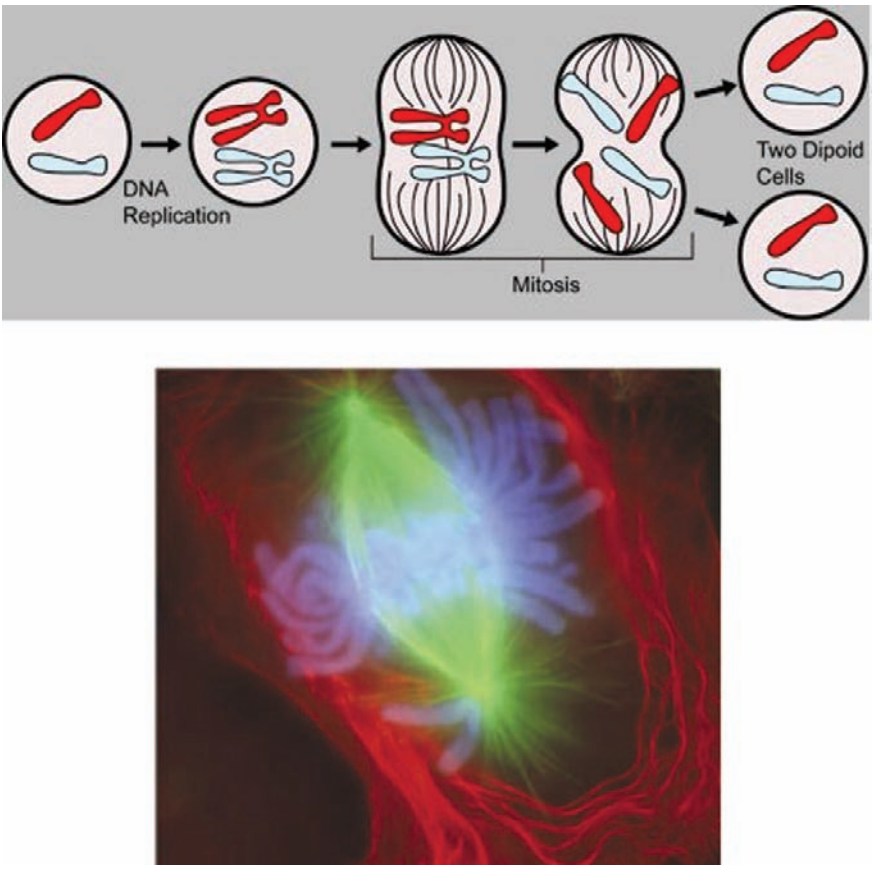

Fig. 2.2 Basic processes during mitosis (Wikimedia Cosmos). Below: an image of a newt lung cell stained with fluorescent dyes undergoing mitosis. The scientists use newt lung cells in their studies because these cells are large, easy to see into, and are biochemically similar to human lung cells (from http://www.nigms.nih.gov/moleculestomeds/images)

and the amount of genetic material is reduced. One parent cell produces four daughter cells. Daughter cells have half the number of chromosomes found in the original parent cell and, with crossing over, are genetically different.

2.2.3.7 Cell Division and Cosmic Catastrophes

Ionizing radiation can damage the DNA and that damage leads to cancer. Radiation produces free OH radicals that can easily break up chemical bonds. Single-strand breaks occur in the DNA. Such free radicals are also produced during normal metabolism and the cell possesses mechanisms to repair such single-strand breaks. The repair makes use of the information redundancy built into the double-stranded DNA molecule and uses the undamaged strand to restore the DNA to its original state. When an entire base becomes separated from the sugar, this is called abasic

site. Furthermore, base alteration can also occur. A double-strand break where two backbones are broken is more severe and cause mutations and cancer.

Cosmic catastrophes that can cause strongly enhanced radiation are as follows:

- Violent solar eruptions (flares, CMEs); they occurred more frequently and with larger intensity during the early phase of the Sun, more than 4 billion years ago.
- Nearby supernova explosions; the probability is not very large but they might cause severe effects (see the relevant chapter).
- Variation of the infalling cosmic ray particle flux due to a weakening of the Earth's magnetic field or the heliosphere. This occurred several times in the evolution of life on Earth.

2.2.4 RNA-Based Life

Ribonucleic acid (RNA) is made from long chains of nucleotides; each of these nucleotides consists of a nitrogenous base, a ribose sugar, and a phosphate. Therefore, RNA is similar to DNA, however, there are several differences: RNA is usually single stranded, RNA contains ribose (DNA contains deoxyribose), RNA contains uracil instead of thymine which is found in DNA, and the chain of RNA is shorter than that of DNA.

The RNA world hypothesis states that RNA-based life could have been a precursor to DNA-based life.The chemical stability of the DNA is greater, however, RNA can still be found, e.g., as rRNA (in ribosomes; catalyzes protein production) in modern cells. Several properties of RNA are important for the beginning of life: ability to self-duplicate, ability to catalyze chemical reactions, and ability to from peptide bonds (e.g., rRNA). RNA has the ability to act as both genes and enzymes.

In chemistry, an enantiomer is one of two stereoisomers that are nonsuperimposable complete mirror images of each other.[17] The term enantiopure means a substance of which only one of two possible enantiomers is present. A theoretical mechanism resulting in the prebiotic appearance of enantiopure ribose, which would be needed for the origin of RNA and the"RNA world" is proposed by Bielski and Tencer [17].

2.3 How Do Organisms Produce Energy?

2.3.1 Oxidation–Reduction

There are two basic processes of energy generation for organisms:

- Photosynthesis,
- Respiration.

[17] For example as one's left and right hands are "the same" but opposite.

Chemically, in both reactions oxidation–reduction reactions are involved. *Oxidation* is the loss of one or more electrons; *reduction* is the gain of one or more electrons, which are added to a compound. In most organic oxidation–reduction reactions, the oxidation of one compound is coupled with the reduction of another compound, catalyzed by an enzyme.

2.3.2 Photosynthesis

Photosynthesis, the transformation of light energy (from the Sun) into chemical energy (sugar), is the principal means of sustaining life at any level. There are only few bacteria that derive their energy from sulfur salts and other inorganic compounds. As a byproduct, oxygen is released. The basic reaction is

$$6CO_2 + 12H_2O + h\nu \rightarrow C_6H_{12}O_6 + 6O_2 + 6H_2O \tag{2.2}$$

Photosynthesis takes place in chloroplasts or in cells with membranes in which chlorophyll is embedded. All the world's plant organisms (in land and in the oceans) produce up to 220×10^9 t of sugar each year. Ninety-four percent of the dry weight substances of green organisms is produced by photosynthesis (amino acids for proteins, etc.). Between 40% and 50% of the oxygen released in the atmosphere originates from photosynthesis of organisms living in the oceans. There are well-known pollutants that can stop the photosynthesis. One example is the PCBs (polychlorinated biphenyls). They are used in electrical insulators and may stop photosynthesis in concentrations as low as 20 ppb. The total amount of sugar produced by photosynthesis in plants determines, finally, how big the population could grow on the Earth.

The principal elements involved in photosynthesis are carbon dioxide, water, light, chlorophyll.

2.3.2.1 Carbon Dioxide

Our atmosphere consists of approximately 78 % nitrogen and about 21% oxygen. Among the remaining gases there is (at present) about 0.036% (360 ppm) carbon dioxide. One acre[18] accumulates more than 2,500 kg of carbon from the atmosphere during a growing season which corresponds to 11 tons of CO_2. The total present atmospheric supply of CO_2 would be completely used up in about 22 years if it were not constantly being replenished. Plant and animal respiration, decomposition, natural fires, and volcanoes replace the CO_2 at roughly the same rate at which it is removed during photosynthesis but human activities (use of fossil fuels, pollution, and deforestation) disrupted this equilibrium. The increase of this greenhouse gas leads to global warming (the global warming since the 1980s cannot be explained only by an increase of the solar luminosity) and by the year 2100, the global temperature will increase between 1.0°C and 3.5°C. An increase in carbon dioxide levels

[18] 1 acre = 4046.85 m^2.

(in 1800, the CO_2 content was 280 ppm[19]; in 2001, it was 360 ppm; the yearly increase is nearly 2 ppm) enhances photosynthesis and increases food production. On the other hand, however, insects, bacteria, and viruses proliferate with warmer temperatures. Fertilizing plants with carbon dioxide in commercial greenhouses has increased yields by more than 20%. Thus, plants have the potential to limit temporarily elevated carbon dioxide levels in the atmosphere. A global temperature increase would perhaps accelerate both plant photosynthesis and animal respiration.

2.3.2.2 Water

Less than 1% of all the water absorbed by plants is used in photosynthesis, the remainder is transpired, incorporated into cytoplasm, etc. Water is used as the source of electrons and the oxygen released is a byproduct and does not come from the carbon dioxide. In case of short supply, photosynthesis becomes limited and the stomata close and the carbon supply is reduced.

2.3.2.3 Light

Forty percent of the radiant energy we receive is in the form of visible light. Leaves absorb about 80% of the visible light. There is a complex connection between light intensities and photosynthesis for different plants. In some plants, photosynthesis will not increase when they receive more than 670 $\mu mol\, m^{-2}$. [20] Plants on a forest floor can survive with less than 2% of full day sunlight, some mosses as low as 0.05% . Most land plants grow naturally at 30% of full day sunlight.

In case of too much sunlight, two processes may occur: photorespriation (uses oxygen and releases CO_2) and photooxidation (bleaching, destruction of chlorophyll). Photooxdiation occurs in autumn, a breakdown of the chlorophyll in leaves.

2.3.2.4 Chlorophyll

All the different types of chlorophyll molecules contain one atom of Mg. They are similar in structure to the heme in hemoglobin.[21] Each molecule has a long lipid tail, which anchors the chlorophyll molecule in the lipid layers of the membranes of cells. *Chlorophyll a* is blue-green in color ($C_{55}H_{72}MgN_4O_5$); *Chlorophyll b* is yellow-green in color ($C_{55}H_{70}MgN_4O_6$). Usually chloroplasts contain three times more chlorophyll a than b. The more the chlorophyll a in the cell, the brighter green the cell. There are also other pigments such as the carotenoids (yellow-orange), phycobilins (blue or red, found in cyanobacteria and red algae), and chlorophyll c, d, and e.

[19] Parts per million.

[20] On a clear summery day at noon the intensity of sunlight is 2000 $\mu mol\, m^{-2}$.

[21] Iron containing red pigment that transports oxygen in the blood.

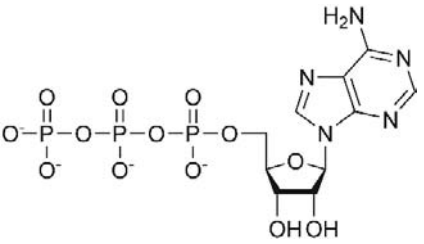

Fig. 2.3 Structure of adenosine-triphosphate (ATP)

Chlorophylls absorb light primarily in the violet to blue and in the red wavelengths. They reflect green wavelengths – therefore, they appear green. Each of the above mentioned pigments has a distinctive pattern of light absorption. In the process of absorption, the energy levels of some of the pigment's electrons are raised.

There are two types of photosynthetic units in most chloroplasts, called photosystem I and photosystem II. Photosystem II was discovered after photosystem I but today we know that the reactions that take place in photosystem II come before those of photosystem I. Both photosystems produce ATP (adenosine triphosphate) (Fig. 2.3) but only organisms that have both can produce NADPH (nicotinamide adenine dinucleotide phosphate) and oxygen as a byproduct. 2.8 billion years ago the cyanobacteria evolved from primitive bacteria and chlorophyll a and photosystem II developed.

The photosystem I consists of more than 200 molecules of chlorophyll a, small amounts of chlorophyll b, carotenoid, and a P_{700} reaction center molecule. Only this molecule can use the energy of the infalling photons; the other molecules act as antenna pigments – they function like an antenna in gathering and passing light energy. The letter *P* stands for pigment and the number for the peak in the absorption spectra of light (photosystem I 700 nm, photosystem II 680 nm).

The photosystem II consists of chlorophyll a, β-carotine,[22] and the reaction center is P_{680}.

These complex mechanisms were invented by the cyanobacteria 2.8 billion years ago when water was already abundant, thus, the process to generate oxygen was facilitated. The oxygen in the Earth's atmosphere increased, therefore, energy-efficient aerobic-respiration-based life could evolve.

Photorespiration occurs when the weather is hot and dry. The stomata are closed and, thus, carbon dioxide is prevented from entering the leaf. Therefore, the concentration of carbon dioxide drops and the concentration of oxygen increases. When the carbon dioxide concentration drops below 50 ppm, rubisco fixes oxygen and photorespiration starts; light is still needed to generate ATP and NADPH. Carbon dioxide is then released.

2.3.2.5 Photosynthesis and Cosmic Catastrophes

Photosynthesis on Earth produced the first signatures of life that can be detected astronomically. Photosynthetic-relevant radiation occurs in the near infrared at bands

[22] Precursor of vitamin A.

0.93–1.1 μm, 1.1–1.4 μm , 1.5–1.8 μm, and 1.8–2.5 μm and plays an important role in the search for extraterrestrial life [122].

As it was shown, photosynthesis is a very complex process depending on many factors; to start the process, the infalling electromagnetic radiation is critical. A change of the infalling light spectrum, therefore, strongly influences the photosynthetic processes. Such a shift in the infalling electromagnetic spectrum can occur by

- The solar output slightly shifted both in luminosity and also in the spectral energy distribution. The early Sun emitted only about 70% of its present output.
- Nearby supernova explosion, an increase of the shortwave energy flux will occur.
- The consequences of an impact are that dust and aerosols will change the transparency of the atmosphere.

The carbon dioxide content is also critical and can be strongly influenced by an impact of an asteroid or by large volcanic eruptions on earth. All these topics will be discussed in more detail.

2.3.3 Respiration

Basically, respiration is the release of energy from glucose molecules that are broken down to individual carbon dioxide molecules. This process takes place in all active cells, regardless of whether or not photosynthesis occurs simultaneously in the cells.

Respiration is initiated in the cytoplasm and completed in the mitochondria. No oxygen is needed to initiate the process, but in aerobic respiration, the process cannot be completed without O_2.

The simplified reaction is as follows:

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O + \text{energy} \tag{2.3}$$

That means

glucose + oxygen → carbon dioxide + water

This aerobic respiration is the most widespread type of respiration. The energy released for this process is 2,880 kJ/mole or 36 ATP.

Anaerobic respiration and fermentation were important when there was no oxygen in the atmosphere (at the early phase of the Earth's evolution). The release of energy is only 6% of the amount that is released by aerobic respiration. Two well-known forms of fermentation are

$$C_6H_{12}O_6 \rightarrow 2C_2H_5OH + 2CO_2 + \text{energy(ATP)} \tag{2.4}$$

Thus, on the right hand side, ethyl alcohol and carbon dioxide are produced. Cells can only survive if the concentration of alcohol is less than 12%. This is well

known from yeasts that ferment sugars. As soon as the alcohol concentration builds up beyond 12%, these cells die and fermentation is stopped.

Another well-known reaction is

$$C_6H_{12}O_6 \rightarrow 2C_3H_6O_3 + \text{energy(ATP)} \tag{2.5}$$

On the right hand side, there occurs lactic acid and energy. The actual amount of energy stored is only 29% of the 200 kJ/mole of energy released approximately in aerobic respiration.[23] The energy released during that anaerobic process is 120 kJ/mole or 2 ATP.

For the strict anaerobes, the presence of oxygen is lethal. An example is *C. tetani* that causes tetanus.

Note also that the terms anaerobic respiration and fermentation are used interchangeably, meaning respiration in an environment of little or no oxygen.

Three factors affect respiration: temperature, water, and oxygen.

Temperature: when temperatures rise from 20°C to 30°C, the respiration rate of plants doubles or even triples. The faster the respiration, the faster energy is released from sugar molecules. In harvested fruits, respiration continues, and, therefore, most fresh fruits are kept under refrigeration to lower respiration rate and water loss and dissipate the heat that is generated by this process. Heating to temperatures > 40°C inactivates most enzymes. Only cyanobacteria (for example) have adapted to hot places above 60°C.

Water: the water content in living cells is more than 90%. In seeds, the water content is reduced to less than 10 % and, also, respiration occurs at extremely low values.

Oxygen: when foods are stored, a reduced oxygen environment helps to lower the respiration rates. In warehouse, where crops are stored, the oxygen content is lowered to 1...3% by replacing it with N. If the oxygen content becomes too low, however, uncontrolled fermentation could start.

2.3.3.1 Respiration and Cosmic Catastrophes

Since respiration depends mainly on temperature, water, and oxygen, we shortly mention possible catastrophes that can influence these factors. The temperature depends on the amount of infalling solar radiation, which, as we have mentioned already, slowly changed during the solar evolution. Besides this slowly varying solar luminosity over the course of billion years, there are also short-term variations due to solar activity. Today these variations are low in amplitude but were by a factor of 100 larger in the case of early Sun. The strongest temperature variations could have been caused by impacts of asteroids. Dust in the atmosphere lead to a global

[23] 1 kcal = 4.184 kJ is the energy required to raise the temperature of 1 kg of water from 14.5 to 15.5°C.

temperature decrease, if the impact occurred in ocean, the amount of additional water vapor, which is an effective greenhouse gas, lead to an increase in temperature.

Also, anaerobic respiration strongly depends on temperature. Anaerobic respiration derived from the period when there was no oxygen in the earth's atmosphere and is also important as adaption in habitats where there is no oxygen. Therefore, in the case of strong reduction of oxygen in an atmosphere, organisms are able to survive a hostile environment. The end-Permian mass extinction occurred during an interval characterized by global warmth and falling atmospheric oxygen levels [165].

The dissolved oxygen content in the world ocean is variable on the decadal timescale [68].

2.3.4 Digestion and Assimilation

As we have shown, sugars are produced by photosynthesis. These undergo many transformations. Some are used directly in respiration; others are transformed into substances like lipids, proteins, and other carbohydrates. Some important carbohydrates are sucrose, starch, and cellulose. By the process of assimilation, these products are used in building the protoplasm and cell walls. The conversion of starch and other carbohydrates to soluble forms is called *digestion*. Let us look at one example—conversion of the disaccharide maltose (malt sugar) in two molecules of glucose with the aid of enzymes:

$$C_{12}H_{22}O_{11} + H_2O \rightarrow 2C_6H_{12}O_6 \tag{2.6}$$

The enzyme involved here is maltase.

Similar to this process, fats are broken down to their component fatty acids and glycerol, and proteins are digested to amino acid building blocks. In animals, special digestive organs were developed. Digestion is similar in both plant cells and animal cells.

2.4 Biological Classification

The goals of biological classification are as follows: assign names to organisms and group them according to their relatedness. If two organisms are related, this means that they share a common evolutionary history or phylogeny, and it follows that both groups evolved from common ancestors. The study of evolutionary relationships is also called *phylogenetics*.

First, organisms were classified into plants (kingdom Plantae) and animals (kingdom Animalia). Today, a system of five kingdoms is accepted (Fig. 2.4). The Kingdom monera: includes the prokaryotes and is divided into Eubacteria (true bacteria) and the Archaea. Just for an illustration, let us give the biological classification of the *Homo sapiens*: Domain: Eukaryota, Kingdom Animalia, Phylum Chordata, Class Mammalia, Order: Primates, Family: Hominidae, Genus: Homo, Species: sapiens.

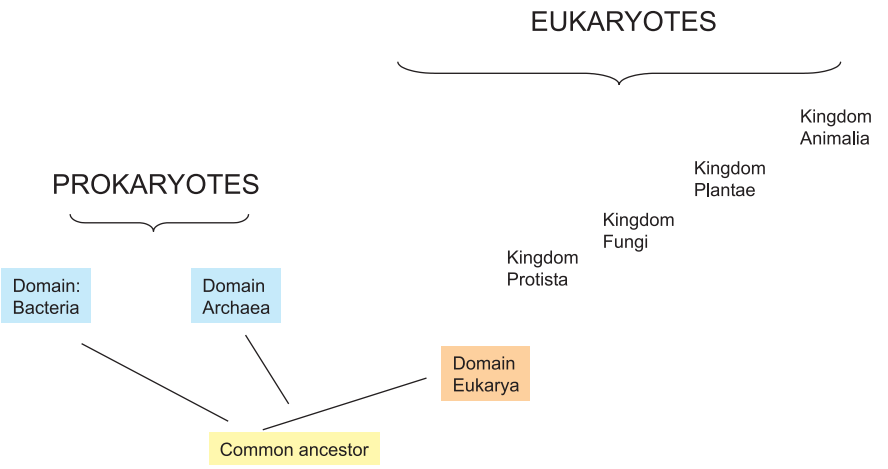

Fig. 2.4 The major groups of living organisms on Earth

2.4.1 Bacteria

The cells are very small, much smaller than single-celled eukaryotes. Under ideal conditions they undergo fission every 10–20 minutes. This leads to an enormous mass: within 36 hours, the number becomes $\sim 2^{100}$. Of course, this is only theoretical and their number is limited by exhaustion of food supplies and accumulation of toxic wastes. Some bacteria have flagella 5–10 micrometers in length. They are used for movement through fluid media.

Bacteria can be divided into heterotrophic and autotrophic. Heterotrophic bacteria cannot synthesize their own food and depend on other organisms. The majority of heterotrophic bacteria are saprobic[24] bacteria. These (as the fungi) are responsible for decay and recycling of organic matter in the soil. Another class of heterotrophic bacteria are parasites — they depend on other living organism for their food.

The autotrophic bacteria are capable of synthesizing organic compounds from simple organic substances. Some autotrophic bacteria like cyanobacterio and chloroxybacteria carry out photosynthesis and oxygen, as a byproduct of these reactions, is set free. The purple sulfur bacteria, the purple nonsulfur bacteria, and the green sulfur bacteria also live from photosynthesis but no oxygen is released. The general reactions here are

$$CO_2 + 2H_2S \rightarrow (CH_2O)n + H_2O + 2S \tag{2.7}$$

The chemoautotrophic bacteria obtain their energy through chemical reactions involving the oxidation of reduced organic groups such as $NH_3, H_2S,$ and Fe^{++}.

Another group of chemoautotrophic bacteria are the iron bacteria. They transform soluble compounds of Fe into insoluble substances that accumulate as deposits,

[24] An organism that derives its nourishment from nonliving or decaying organic matter.

for example, in water pipes. The sulfur bacteria convert hydrogen sulfide gas to elemental sulfur and sulfur to sulfate.

Since bacteria are ubiquitous, any accumulation of household garbage, leaves, weeds, etc., heaped together will become decomposed and heat is generated (up to 70°C). Due to the increase of heat, organisms that are adapted to higher temperatures replace the others. Also, many weed seeds and most disease organisms are killed during this process.

The Bubonic plague (Black Death) are bacteria transmitted by fleas, deerflies, and ticks of lice that have been parasitizing infected animals (particularly rodents). In 1665, in London, hundreds of thousands perished from the Black Death. Between 1347 and 1349, 1/4 of the entire population of Europe (25 million persons) were killed. It is still present in some parts of Asia, and even in Europe and South America. It is possible to control this plague by controlling rats and fleas and the use of vaccines.[25]

Many harmful pests and weeds can be significantly limited through the use of biological controls. The *Bacillus thuringiensis* (often referred to as Bt) is effective against a wide range of caterpillars and worms, like the peach tree borers, European corn borers, cabbage worms, etc.

2.4.1.1 Cyanobacteria and Prochlorobacteria

There is a difference between true algae and blue-green algae, the cyanobacteria. The cyanobacteria have prokaryotic cells and are much more like true bacteria. The main differences between the cyanobacteria and bacteria are

- Cyanobacteria have chlorophyll which is found in higher plants,
- Cyanobacteria have blue green phycocyanin and red phycoerythrin pigments,
- Cyanobacteria can both fix nitrogen and produce oxygen (normally nitrogen-fixation is essentially an anaerobic process).
- Also, the prochlorobacteria are capable to carry out photosynthesis, however, they do not produce oxygen.

The cyanobacteria are found in pools, particulary if the water is polluted, in fresh and marine waters, and even in ice-covered lakes of the Antarctic. In the oceans, cyanobacteria are the principal photosynthetic organisms in plankton (the cyanobacterium Synechococcus at a size of 0.8–1.5 μm occurs in concentrations 10,000 cells per milliliter). A species is also found in the different temperature ranges of Yellowstone National Park at water temperatures up to 85°C. There, the bacteria precipitate carbonate deposits – a rocklike substance called travertine (2 to 4 mm/week).

After volcanic eruptions, cyanobacteria are often found to be the first photosynthetic organisms on bare lava. They are even found on the shells of turtles and snails. Others live symbiotically. Cells of cyanobacteria occur in colonies- often in chains or filaments, sometimes branched. The individual cells are held together by

[25] Immunity lasts for about 6 to 12 months.

gelatinous sheaths, which can be colored red, brown, green blue, and even violet. There are cyanobacteria that occur symbiotically – they lack cell walls and function as chloroplasts. There is a speculation that chloroplasts originated as cyanobacteria or prochlorobacteria that live within other cells. That assumption is also supported by the fact that when eukaryotic cell containing chloroplasts divide, the chloroplasts divide at the same time.

Microorganisms 3 billion years old from the Precambrian of South Africa were found by Barghoorn and Schopf [10]. In 1993, Schopf [218] described fossils that resemble cyanobacterium-like microorganisms and the fossils are about 3.5 billion years old. However, these fossils are no longer considered to be morphologically similar to cyanobacteria. It is assumed that about 3 billion years ago, these organisms began producing oxygen, which started to accumulate in the earth's atmosphere. It lasted, however, 2 billion years before oxygen became a substantial part of our atmosphere. The ozone layer became an effective shield against shortwave harmful solar UV radiation 1 billion years ago. Before that time, the organisms were only protected by their watery environment. Again, the important role of the cyanobacteria for the entire history of the evolution of life on Earth is seen. In water, they are often at the bottom of the food chains.

The Prochlorobacteria were detected in 1976[26] They have chlorophylls a and b like the plants but no phycobilin pigments that are typical for the cyanobacteria.

2.4.1.2 Archaebacteria

There are three distinct groups of archaebacteria.

1. Methane Bacteria: they are killed by oxygen and survive only under anaerobic conditions found in swamps, oceans, lake sediments, hot springs, animal intestine tracts, and other environments. Methane bacteria generate energy through formation of methane gas (CH_4) from carbon dioxide and hydrogen. Methane will burn when it constitutes only 5% to 6% of the air (dancing light over swamps seen at night). To our civilization, methane could be extremely useful. It is the principal component of natural gas. It can be estimated that 9 kg of horse manure (or 4.5 kg of pig manure) fed daily into a methane digester will produce all the gas needed for the average adult's cooking needs in northern latitudes. Methane is clean, practically nonpolluting, safe and nontoxic and can be used as a motor fuel.
2. Salt Bacteria: observed from above, salt evaporation ponds and other salt containing shallow bodies of water often have a red appearance. This is caused by the salt bacteria. They carry out photosynthesis with the aid of the red pigment rhodopsin.
3. Sulfolobus Bacteria: they occur in sulfur hot springs and are thermophilic bacteria. Their environment is unusually hot (up to 110°C) and exceptionally acidic (pH = 2).

[26] R.A. Lewin.

2.4.2 Viruses, Viroids, and Prions

2.4.2.1 Viruses

During the time of L. Pasteur,[27] all infectious agents including bacteria, protozoans, and yeast were called viruses. One of Pasteur's collaborators discovered that porcelain filters would block out bacteria but not smaller elements. Agents that penetrated such porcelain filter were called filterable viruses or just viruses. Viruses cause many well known diseases like smallpox, measles, mumps, chicken pox, polio, yellow fever, influenza, fever blisters, and the common cold.

How are viruses defined? First, they have no cellular structure. In 1946, W. Stanley got a Nobel Prize because he demonstrated that a virus causing tobacco mosaic (plant disease) could be isolated, purified, and crystalized. These crystals can be stored and always still produce the disease in the plants. The second property of viruses is that they do not grow by increasing their size or dividing. Furthermore, they do not respond to external stimuli, cannot move, and there is no independent metabolism. They are about the size of large molecules (15–300 nm). A typical teaspoon of sea water contains more than 10^9 viruses.

Viruses consist of a nucleic acid core surrounded by a protein coat. The nucleic acid core contains either of DNA or RNA–never both. The protein coats are complex (many have 20 sides), some have a distinguishable head and tail. Viruses that attack bacteria are called *bacteriophages* (sometimes only the word phage is used).

Typical for viruses is their reproduction. They attach to a host cell. Then, they penetrate to the interior. Their DNA or RNA directs the synthesis of new virus molecules, which are then released from the host cell. Usually, the host cell dies. Some viruses can mutate very rapidly, e.g., influenza. They are able to attack organisms that previously had been immune to them. New vaccines must be developed. Many forms of cancer (abnormal cell growth) are caused by viruses. Cells of higher animals that are invaded by viruses produce a protein–*interferon*. This cause the cells to produce a protective protein, which prevents the propagation of many types of viruses. It also inhibits viruses from causing tumors. Thus, interferon can be used for controlling many forms of cancer.

2.4.2.2 Viroids and Prions

Many plant diseases are caused by viroids (e.g., potato spindle-tuber disease). Viroids consist of circular strands of RNA and occur in the nuclei of infected plant cells. They may remain latent in some plants and cause severe symptoms in others. Viroids are transmitted from plant to plant via pollen.

Prions are also smaller than viruses. They are particles of protein that cause diseases of animals and humans.

---

[27] 1822–1895.

2.4.3 The Kingdom Protista

As it was pointed out, less than 1 billion years ago, all living organisms on Earth lived in the oceans. Such an environment was necessary because of the following reasons:

- Protection from drying out.
- Protection from intense solar UV radiation.
- Protection form large temperature fluctuations.

About 400 million years ago, green algae — members of the kingdom protista — started to make the transition from water to land. Thus, plants evolved. An ancestor of today's land plants could be *Coleochaete* — a green freshwater alga that grows epiphyte (i.e., it attaches itself in a nonparasitic manner to another living organism). We find, for example, the production of lignin-like compounds that are important to strengthen the cell walls.

All members of the Kingdom Protista have eukaryotic cells. Algae can be found in Seaweeds, fish-tank films, in swimming pools, etc. The green algae are an important part of the plankton (free floating microscopic organisms). Like the higher plants, green algae store their energy in the form of starch within the chloroplasts. The Chlamydomonas is unicellular and has a pair of whip-like flagella. The *volvox* forms a hollow spherical colony of cells that spin on their axes as the flagella of each cell beats in a coordinated way.

2.4.4 The Evolution of Complex Life, Cosmic Catastrophes

Complex organisms evolved from simple unicellular organisms, some cells became specialized to specific tasks. Complex organisms need more energy to maintain their function, therefore, a transition from photosynthesis to aerobic respiration was necessary. The prize for these evolutionary quantum leaps was a greater vulnerability to changing environmental conditions that result from cosmic catastrophes.

We can state that

- primitive life (bacteria, archaea): less vulnerable to cosmic catastrophes; once having formed on a planet, it is highly probable to survive a cosmic catastrophe.
- complex life: multicellular; more vulnerable to catastrophes; chances to survive cosmic catastrophes are smaller. Needs a stable continuous habitable zone for a long time; cosmic catastrophes led to mass extinction and provided the chance to mutations that evolved to more complexity.

Chapter 3
Stars and Galaxies

This chapter gives an overview of the cosmic environment of the Earth, the stars and the galaxy. It is important to know their evolution because this contributes to the cosmic circle of matter: stars are born out of interstellar clouds and at the end of their evolution, they enrich the interstellar medium with elements heavier than He. These elements have been produced by nuclear fusion in the stellar interior and are important for the origin of planetary systems and the origin of life itself.

The knowledge of stellar evolution is essential for the timescale over which life evolves. If higher lifeforms, as we know from Earth, need several billion of years to evolve, then we must search for a star that does not change in luminosity over such a long time span in order to guarantee a continuous habitable zone.

3.1 Stars: Evolution and Formation

3.1.1 Physical Parameters of Stars

In the case that the host star provides the energy source for life, the habitability on planets around that star strongly depends on stellar physical parameters (see also Sect. 9.2.3) like (i) stellar mass, (ii) stellar temperature, (iii) age of the star, and (iv) stellar activity and others.

There are relations between these parameters. The activity of a star depends on its rotation and magnetic fields. The lifetime of a star depends strongly on its mass. A summary of stellar parameters and their importance for HZ around the stars is given in Table 3.1.

3.1.2 Spectral and Luminosity Classes

The radiation of a star can be analyzed using a spectrograph. According to their spectra, stars can be classified in the following sequence: O-B-A-F-G-K-M. This

A. Hanslmeier, *Habitability and Cosmic Catastrophes*, Advances in Astrobiology and Biogeophysics, DOI 10.1007/978-3-540-76945-3_3,

Table 3.1 Stellar parameters and their influence on habitable zones

Parameter	Influence
Age	Important for total lifetime of star
Mass	Determines the lifetime; stars with high masses evolve too fast
Temperature	Coupled to the mass
Composition	Nearly identical for all stars
Magnetic field	Important; determines stellar activity
Rotation	Important; determines stellar activity

is a sequence of temperature (see Table 3.3): O stars are the hottest, M stars the coolest; the number of absorption lines increase from O to M. Some characteristics are given in Table 3.2. The fact that there appear metal lines in cool stars does not imply that they consist of metals. It is only because of the lower temperature that such spectral lines can be formed; the chemical composition is roughly the same for all types of stars (mainly H and He).

The luminosity of a star depends on (a) temperature $\sim T^4$ and (b) surface which is $\sim R^2$. Since, e.g., a K star may be a dwarf main sequence star or a giant, luminosity classes have been introduced. Class I contains the most luminous supergiants and class II the less luminous supergiants. Class III are the normal giants, class IV the sub giants, and class V the main sequence.

Now, we understand the spectral classification of our Sun: it is a G2V star.

As it will be shown in the chapter about stellar evolution, the lifetime of a star is mainly determined by its main sequence lifetime. Thus, for our investigation about habitability, we can definitely exclude all stars that belong to luminosity classes I–IV because their lifetime is too short (several 10^6–10^8 yr), that means their evolution is too fast, these classes reflect the evolution after the main sequence.

Table 3.2 Spectral classification of stars

Type	Characteristic of spectrum
O	Ionized He, ionized metals
B	Neutral He, H stronger
A	Balmer lines of H dominate
F	H becomes weaker, neutral and singly ionized metals
G	Singly ionized Ca, H weaker, neutral metals
K	Neutral metals molecular bands appear
M	TiO, neutral metals
R,N	CN, CH, neutral metals
S	Zirconium oxide, neutral metals

Table 3.3 Effective Temperature as a function of spectral type

Spectral Type	O	B0	A0	F0	G0	K0	M0	M5
T_{eff} [K]	50 000	25 000	11 000	7 600	6 000	5 100	3 600	3 000

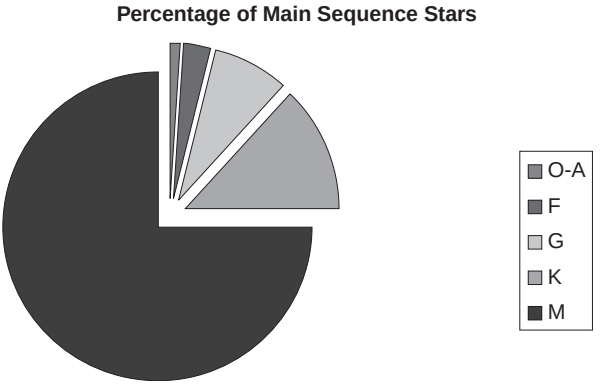

Fig. 3.1 Percentage of spectral classes of main sequence stars

The percentage of the main sequence stars for different spectral types is given in Fig. 3.1. Note that solar-like stars of spectral class G consist only about 8% of all main sequence stars.

3.1.3 Main Sequence Lifetime

To study the habitability in a planetary system around main sequence stars, we must take into account the evolution of the host star. About 1 billion years after the formation of the Earth, the first life forms occurred. Maybe there was life even before, but this is uncertain. So, the host star must be stable for at least over $\sim 10^9$ years for life to evolve.

The evolution and lifetime of a star depends primarily on its mass. The larger the mass, the larger its luminosity, the more the energy needed to provide that luminosity, and the faster its evolution. From these considerations, we can estimate the main sequence lifetime of a star:

$$t_*/\mathrm{t}_\odot = (M_*/\mathrm{M}_\odot)/(L_*/\mathrm{L}_\odot) = (M_*/\mathrm{M}_\odot)^{-2.3} \tag{3.1}$$

which gives the lifetime of a star as a fraction of the Sun's expected lifetime (10^{10} yr). A simplified version of this formula is

$$\tau_* \sim 10^{10} \cdot \left[\frac{M_\odot}{M}\right]^{2.5} \text{ yr} \tag{3.2}$$

In Fig. 3.2, the main sequence lifetime is shown for stars with different masses (in units of the Sun's mass). We see that for stars with masses $> 2\,\mathrm{M}_\odot$, the main sequence lifetime drops below 1 Gyr, thus, even if habitable planets around them exist, life would probably not have enough time to evolve.

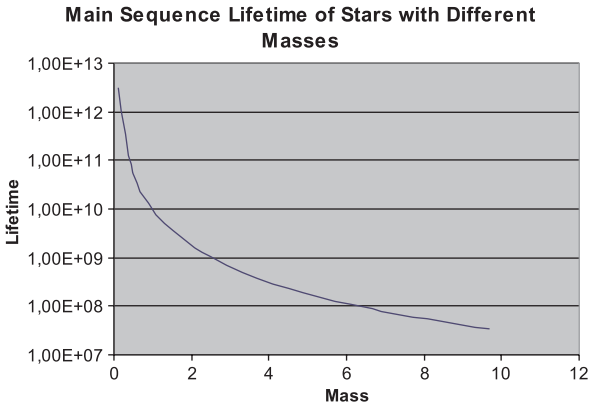

Fig. 3.2 Main sequence stars with different masses and their lifetime

3.1.4 Stellar Evolution

3.1.4.1 The Hertzsprung-Russel Diagram

When the temperature of a star is plotted versus its brightness (corrected for its distance), then stars occur at a defined position in the so-called Hertzsprung-Russel diagram (HRD; see Sect. 9.2.4). If this is done for many stars, it is seen that stars are not randomly distributed in the HRD as illustrated in Fig. 3.3:

- Main sequence stars: most stars are found along a diagonal from the upper left (hot) to the lower right (cool).
- Giants, supergiants: they have the same temperature as the corresponding main sequence stars but are much brighter and must have larger diameters (see Eq. (9.18)).
- White dwarfs are faint but very hot objects; from their location at the lower left in the HRD, it follows that they must be very compact (about 1/100 the size of the Sun).

Considering a large sample of stars, it is evident that we observe young objects and old objects. The fact that most of the stars are found on the main sequence in the HRD is because this phase of stellar evolution lasts longer than all other phases.

In Fig. 3.3 also, the region is indicated where stars with habitable planets have to be expected. These are main sequence stars ranging from temperatures of about 7,000 K to 4,000 K.

3.1.4.2 Evolution of the Sun

We can even plot the evolution of a star like our Sun in the HRD. The main steps in the evolution of the Sun are (compare with Fig. 9.13)

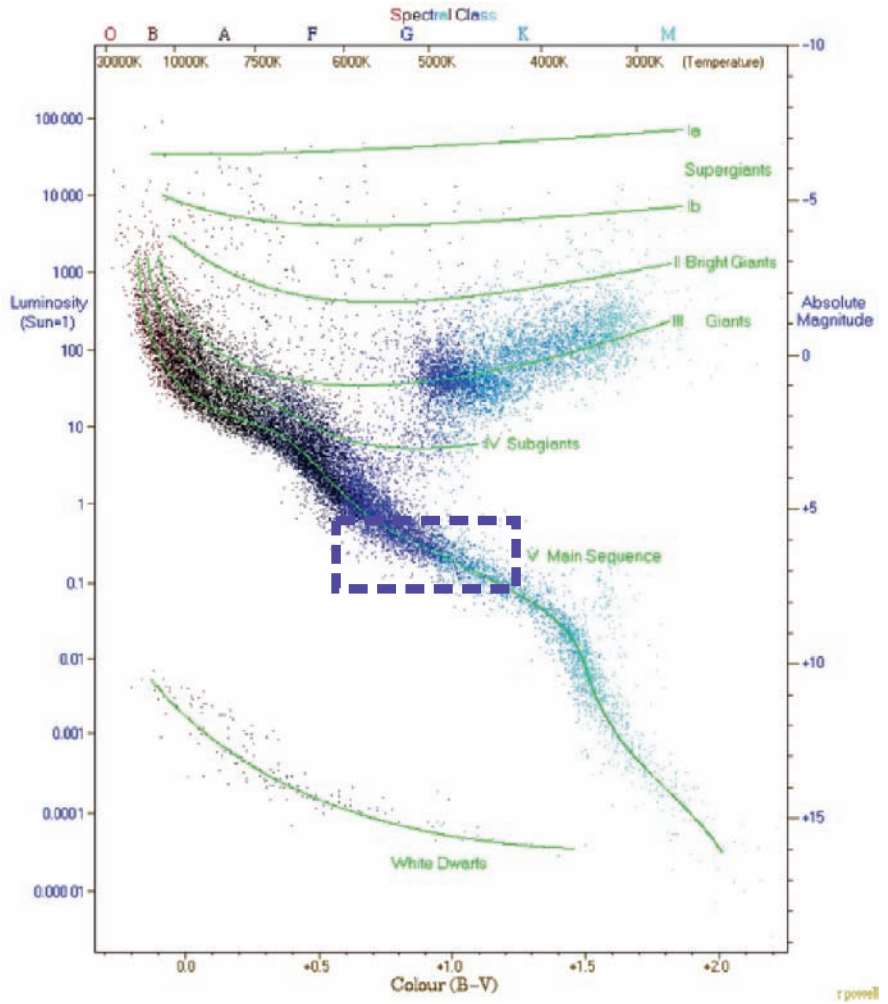

Fig. 3.3 The Hertzsprung-Russel diagram. More than 20,000 stars are plotted and the main sequence stars form a diagonal. The region where solar-like stars are found that could possibly host habitable planets is indicated. Adapted from R. Powell

- Pre-main sequence evolution: the Sun was formed from a protostellar gas and dust cloud and before it has reached the main sequence, where it spends most of its life, the contracting Sun has passed a violent youth, the T Tauri phase.
- As a main sequence star, the Sun changes extremely slowly, remaining there about 10^{10} years. In the core, H is transformed into He by nuclear fusion.
- The Sun evolves to a red giant, it will expand, and the Earth will become part of the solar atmosphere. The expansion starts when all H is transformed to He in the core. Then, a H burning shell supplies the energy. As soon as, in the center, He

burning sets in there occurs a He flash. Then, the Sun will evolve to a red giant for some 10^8 yr, extending beyond the Earth's orbit.
- Finally, the Sun becomes a white dwarf that slowly cools.

During its evolution, the Sun dramatically changes its radius (the subscript $_\odot$ denotes the present day value):

$$1\,\mathrm{R}_\odot(\text{present Sun}) \to\sim 10^3\,\mathrm{R}_\odot(\text{red giant}), \to 0.01\mathrm{R}_\odot(\text{white dwarf}).$$

3.1.4.3 Final Stages of Stellar Evolution

For the main sequence stars, there exists a relation between their mass M and luminosity L:

$$L \sim M^{3.5} \tag{3.3}$$

From (3.3), we see that more massive stars are very luminous and, therefore, they use up their nuclear fuel much more rapidly than low massive stars like our Sun. Massive main sequence stars that are observed today must have been formed in very recent astronomical history.[1]

The final stages of stellar evolution depend on their masses:

- Stars with masses $< 1.4\,\mathrm{M}_\odot$ will end as compact *white dwarfs*, they will reach the size of the Earth and cool down without producing energy in their interiors. Because of their small surface area, this cooling process lasts very long. Concerning planets, so far, no objects were found in the vicinity of white dwarfs. From the described scenario of their evolution, it is probable that planetary systems can survive the transition from a normal star to a red giant and white dwarf. The red giant phase, however, most probably destroys habitability and the white dwarf is very hot and compact, thus, HZ should be near them, but because of their high temperature, these objects radiate strongly in the UV which again destroys biologic activity.
  Zinnecker et al. [276] made a search for planetary companions of several Jupiter masses around known white dwarfs in the Hyades cluster which is at a distance of 45 pc from us (1 pc = 3.26 Light years). These white dwarfs had progenitor masses of about 3 solar masses and are about 625 million years old. Since planets radiate more strongly in the infrared (because of their low temperature), the infrared NICMOS camera at the Hubble space telescope was used for the search, however, the result was negative. Mullally and Winget [172] found evidence for a possible planetary companion of about 2 Jupiter masses at a distance of 2 AU from a white dwarf.

[1] In some large interstellar nebulae, one observes stars that have an age of some 10^5 years.

- Stars with masses $1.4\,M_\odot < M_* < 4-5\,M_\odot$ will end as *neutron stars*. Neutron stars are the result of a supernova explosion. It can be shown by quantum mechanics that for stellar masses larger than $1.4\,M_\odot$, no stable white dwarfs can exist for which the pressure of degenerate electrons stop further collapsing. In stars with larger masses than this limit (which is also called *Chandrasekhar limit)*, the electrons combine with the protons to form neutrons and a neutron star is the result. The radius of these extreme compact objects is about 10 km. Because of spin and magnetic field conservation, they rotate very rapidly and particles are accelerated along their magnetic field lines. Then, this emission can be seen as pulses (the objects are also known as pulsars) if the axes of rotation and magnetic field do not coincide.
  We will discuss in Sect. 4.3.4 that planetary systems around neutron stars were detected, which was a big surprise. Because of their bundled radiation, HZ around neutron stars are very unlikely.
- *Black hole*: when the mass M_* exceeds about 4–5 $M_\odot$, then, the object collapses to a black hole. No matter, no radiation that enters beyond the *Schwarzschild radius* R_S can escape from a black hole.
  The Schwarzschild radius can be easily calculated. The escape velocity from an object with mass M is

$$v_{\text{esc}} = \sqrt{2GM/R} \tag{3.4}$$

  Now, let $v_{\text{esc}} = c$, the speed of light and the radius of such an object would become R_S:

$$R_S = \frac{2GM}{c^2} \tag{3.5}$$

  This is also called the *event horizon* from which nothing can escape. According to general relativity, a black hole's mass is entirely compressed into a region with zero volume, the curvature of space time is infinite. The interstellar matter contains a few atoms per cm^3 and this matter is attracted toward the black hole. The gas starts to form a disk in which the matter is compressed to a high density. In the disk, pressure and friction occurs that cause the disk to glow very bright. Here, enormous amounts of energy are generated. From these facts, it is clear that black holes are extremely hostile to habitability, however, they produce enormous amounts of energy.

The evolution of life is a process that seems to take at least several 10^8 years. The end stages of stelar evolution are not candidates for possible HZ around them. Only in the case of neutron stars (pulsars), planetary systems have been found so far.

3.2 Stellar Evolution and Habitability

3.2.1 Main Sequence Stars

We consider stellar temperatures, masses, and luminosities and the possibility of HZ around them. In Table 3.4, we give the following parameters: spectral class of

Table 3.4 Habitable zones and some stellar parameters

Spectral Class	T_{eff} (K)	Life (y)	Abundance (%)	HZ (AU)
O6V	41 0000	10^6	4×10^{-5}	450–900
B5V	15 400	8×10^7	0.1	20–40
A5V	8200	10^9	0.7	2.6–5.2
F5V	6400	4×10^9	4	1.3–2.5
G5V	5800	2×10^{10}	9	0.7–1.4
K5V	4400	7×10^{10}	14	0.3–0.5
M5V	3200	3×10^{11}	72	0.07–0.15

the central star, its surface temperature, the main sequence lifetime for the star, the percentage of this type of stars in the overall stellar population of a typical galaxy, and the theoretical range of the HZ.[2] Only main sequence stars were considered because stars are found for most of their lifetime on the main sequence. It must be stressed that this is a very rough estimation of HZ because only the flux was taken into consideration and not the wavelength. Because of their high temperatures, the maximum emission for O- and B-stars is in the UV and, therefore, apart from their extreme short lifetimes, their circumstellar environment is expected to be quite hostile. Even the environment of A5V stars is hostile because of high UV radiation and expected lifetime below 10^9 y. In our solar system, we expect that the first primitive forms of life appeared on Earth about 1 Gy after its formation. Our Sun is a G2V star and its total lifetime on the main sequence will be about 9 Gy.

From Table 3.4, we see that stars of spectral type F5V have a main sequence lifetime of 4×10^9 yr and their HZ are between 1.3 and 2.5 AU because of their larger temperature. Therefore, only stars of spectral type later than F5V seem to have large enough main sequence lifetime so that higher life could evolve. Of course, these considerations are made under the assumptions that life everywhere in the universe evolves or has evolved at the same rate and speed as on Earth. Nobody is able to prove or disapprove this hypothesis.

Considering stars that are cooler than the Sun, we first notice a strong increase in their lifetime. Thus, life would have longer time spans to evolve in such planetary systems. However, as it will be discussed in detail later, there are also some limitations for such hypothetical planetary systems.

3.2.2 Habitable Zones and Stellar Evolution

During their evolution, stars become more luminous. This can be seen by a simple consideration. On the main sequence, stars are in hydrostatic equilibrium; the gravitational force equals the pressure force. The latter is given by

[2] Adapted from Ref. [249].

$$P \sim \sqrt{\frac{R_g T}{\mu}} \tag{3.6}$$

R_g is the gas constant, T the temperature, and μ the mean molecular weight. The gas pressure depends on T and μ. Because of the fusion of hydrogen to helium, the mean molecular weight μ increases, which would decrease the pressure. This is compensated by an increase in temperature and, therefore, the luminosity of the star increases.

The past evolution of solar luminosity can be represented by

$$L_\odot^{\text{past}}(t) = \left[1 + \frac{2}{5}\left(1 - \frac{t}{t_\odot}\right)\right]^{-1} L_\odot^{\text{pres}} \tag{3.7}$$

This formula was given by Gough [78]. $L_\odot^{\text{pres}}$ is the luminosity of the Sun at present time $t_\odot \sim 4.6 \times 10^9$ yr.

For the future evolution, the formula given by Turck-Chieze et al. [248] can be applied:

$$L_\odot^{\text{fut}} = \left[5.59\,\text{Gyr}\,\frac{1}{t} - 1.39 + 0.26\,\text{Gyr}^{-1}\,t\right] L_\odot^{\text{pres}} \tag{3.8}$$

and both curves are merged in Fig. 3.4.

Generally, HZ around F stars are larger and further away from the star than for the Sun; HZ around K and M stars are smaller and nearer to the star. It is interesting to note that the widths of all these HZs are nearly the same if the distance is expressed on a logarithmic scale (see also Fig. 1.1).

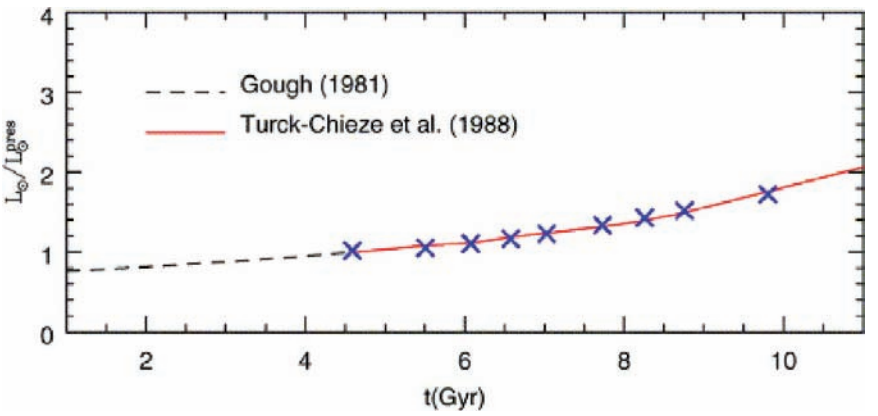

Fig. 3.4 The evolution of the solar luminosity

3.3 Galaxies

The word galaxy is derived from the Greek word *Galaxias* meaning milk. On a clear, dark summer night, the Milky Way can be observed as a hazy band across the celestial sphere. Already, in ancient Greek, it was supposed that it is composed of billions of stars (*Democrit*, 450–370 BC).

Stars are not isolated objects in the universe but they form huge aggregations called *galaxies*. Our Galaxy, the *Milky Way* Galaxy, is the host of the solar system together with at least 200 billion other stars, many of them are supposed to have planetary systems. In this section, we will first discuss the basic properties of our own system and then discuss other galaxies. Besides the host star, also the galactic environment is crucial for the stability of HZ around the stars.

3.3.1 Our Galaxy

The Galaxy comprises at least 200 billion stars including the solar system, gaseous nebulae, dark dust clouds, and the so-called dark matter that can only be detected by its gravity and does not seem to emit electromagnetic radiation. The dark matter most probably exceeds the luminous matter by a factor of 10–20.

Some information about the structure of the Galaxy can be obtained by very simple observations. The galaxy is the brightest in the constellation *Sagittarius*, therefore, in this direction, the galactic center can be seen (however, there are many dark interstellar clouds in this area so that it can not be observed in the visible range of the spectrum). The solar system lies close to the galactic plane. This follows from the fact that the Milky Way divides the night sky into two roughly equal hemispheres.

Another source of information about the position of the solar system is to count extragalactic objects, galaxies. One finds that there is a *zone of avoidance* about the galactic plane, where galaxies become very rare because galactic interstellar matter, dark clouds of dust, and gas absorb the light of these objects at low galactic latitudes.

Galaxy: Basic facts Diameter: 100,000 Light years
Mass: $4.8\times10^{11}\,\mathrm{M}_{\odot}$; this corresponds to 200...400$\times10^{9}$ stars
Age: $13.6\pm0.81\times10^{9}$ years (from the age of globular clusters)
Hubble Type: SBc (loosely wound barred spiral)

The appearance of our Galaxy from outside would be that of a huge spiral system (at least four spiral arms). From the edge, it would be seen as a very thin disk with a central bulge. The solar system is located approximately 30,0000 Lyr from the center near the galactic equator (Fig. 3.5). The position of the Sun and the solar system in the galaxy is as follows:

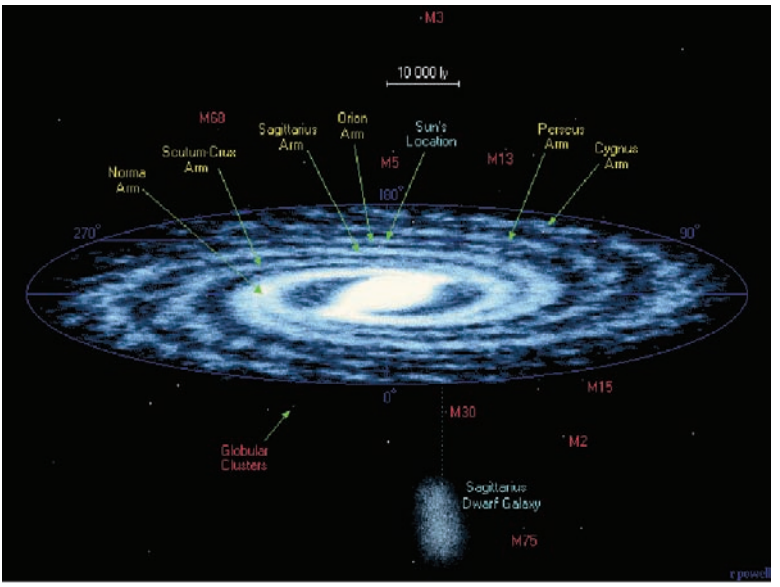

Fig. 3.5 Sketch of the Galaxy

```
Solar System and Galaxy:
Distance from galactic center: 8,600 pc
Distance from galactic plane: 15 pc
```

Let us consider a simple model: assume that the galaxy were reduced to 130 km, then, the size of the solar system would be 2 mm!

3.3.1.1 Main Components

Stars can also be classified into several *populations*. Population II stars are metal poor stars.[3] Population I stars contain more metals. Population I stars are found mainly in the spiral arms, whereas the population II stars can be found in the bulge as well as in the halo. The distinction into metal poor and metal rich stars is very important for the topic of habitability. Metal poor stars were formed from metal poor gas clouds and the probability to find planetary systems there is very low, since there were no elements like Fe, Ni, Si, C, etc., to form terrestrial, like planets.

The main components of our galaxy are then

- Disk: approximately 300 pc thick and 30 kpc (about 100,000 Lyr) in diameter. It is made up predominantly of population I stars. These stars are blue and relatively young: between $\sim 10^7$ and $\sim 10^9$ years.

[3] The term metal in astrophysics denotes any element that is heavier than He.

- Bulge: about the center of the galaxy, there is a flattened spheroid of dimension 1 kpc $\times$ 6 kpc. In this region, the stellar density is high the stars are old and red. The oldest stars exceed 10^9 years.
- Center: as it has been stated above, the center itself cannot be observed in visible light because of absorption by interstellar matter. Radio observations can be used as there is no absorption at these wavelengths. From the dynamics of stars near the center, it became evident that there must be a supermassive black hole (SMBH).
- Halo: in the galactic halo, we find globular clusters which are spherical in shape and contain 10^4 to 10^6 stars that are very old and metal poor. There is also evidence that dark matter is found there. This follows from observations of rotational velocities of objects as a function of their distance to the galactic center.

3.3.1.2 Rotation of the Galaxy

One of the most important properties of galaxies is their rotation. The *Sun* (and, thus, the solar system) moves about the center of the galaxy, the orbital speed is 217 km/s, and it takes $225–250\times10^6$ years for one revolution. Therefore, during the existence of the solar system and the Earth, between 20 and 25 revolutions have occurred.

The solar apex, the direction in which the Sun travels through space, lies in the constellation *Hercules* (near *Vega*). The Sun's orbit around the galaxy is elliptical with additional perturbations due to the galactic spiral arms and nonuniform mass distributions.

Using Kepler's third law, from the rotation of a galaxy, one can derive its mass. One would expect that the rotation velocity decreases as one moves further away from the center. However, this is not the case, on the contrary, the rotation velocity even increases for stars that are at larger distances from the galactic center than the Sun. This can only be explained by the presence of *dark matter*, since we cannot observe radiation from this matter in any wavelength region.

What is the nature of dark matter? Dark matter is matter that does not emit electromagnetic radiation; its presence can be inferred only from gravitational interaction. The objects that make up dark matter are unknown but there are several candidates:

- MACHOs: massive compact halo objects. These objects could be neutron stars, black holes, black dwarfs, brown dwarfs, dwarf planets.
- WIMPs: weakly interacting massive particles. These particles should interact only through the weak force and gravity, therefore, they can not be "seen." WIMPs must have been created during the Big Bang and be very massive.

3.3.1.3 The Spiral Structure

Galaxies rotate and the spiral structure of a galaxy must be maintained by some mechanism.

Density waves are triggered by the distribution of the stars but also by encounters of the Galaxy with neighboring galaxies and may be the cause for the spiral arms. Density waves are very important for the compression of interstellar matter. Through density waves, the dust clouds become gravitationally unstable, collapse, and starforming regions emerge. Therefore, the spiral arms are the regions where young stars, often very hot, massive, and shortlived, are found. The most current density wave can be seen in the H II regions (that are regions in which hydrogen is ionized), the elder next in blueish clouds and associations of hot young stars, and then, the next in the distribution of type II supernova remnants. Older stars have dissipated into the more diffuse background and are also called intermediate population I stars.

The solar system is located in the Orion arm (the arms are named after the constellations where substantial parts of them are situated). The other arms are (from outside to inside)

- +II, outer arm
- +I, Perseus arm
- O, Local arm, Orion arm
- -I, Sagittarius arm
- -II Scutum arm
- Norma arm

The arms can be traced by different objects: diffuse nebulae, open star clusters, associations, supernova remnants, etc.

3.3.2 Galaxies

3.3.2.1 Types of Galaxies

Until the observations in the 1920s made by *E. Hubble*, it was unclear whether there exist other galaxies (see examples Figs. 3.6, 3.7). Hubble was able to determine the distance of the nearest galaxies[4] by resolving their outer parts into individual stars. The nearest galaxy is the Andromeda galaxy, M31, (see the original paper given by E. Hubble [110]) at a distance of more than 2 million light years. Like our system it appears as a spiral galaxy. However, there exist other types of galaxies and a classification scheme still often used is that of Hubble:

- Elliptical galaxies; diameters between 1 and 50 kpc,
- Spheroids,
- Spirals; diameters between 10 and 30 kpc,
  - Normal spiral,
  - Barred spiral,
- Irregular galaxies; diameters between 5 and 20 kpc.

[4] An overview was given recently by: [232].

Fig. 3.6 A barred Galaxy NGC 1300. Photo credit: NASA/ESA Hubble Space Telescope

The largest galaxies are giant elliptical galaxies. It is believed that elliptical galaxies form due to the interaction of galaxies by merging or collisions. Giant ellipticals are often found near the core of large galaxy clusters.

Masses of galaxies can be deduced from Kepler's third law or from the Virial theorem. The largest galaxy, M87, in the center of the Virgo cluster has a mass of $3 \times 10^{13}\,\mathrm{M}_{\odot}$.

Fig. 3.7 The so called Antenna Galaxies, an example of colliding galaxies. Photo credit: NASA/ESA Hubble Space Telescope

Neutral hydrogen can be found by measuring the 21 cm (or 1,420 MHz) H line (this line occurs from a transition of nuclear spin parallel or antiparallel to electron spin in the hydrogen atom). It follows that neutral H (also called H I) increases from type S0-Sa-Sb-Sc-Irr from 1% to 25% of the total mass of the galaxy.

3.3.2.2 Supermassive Black Holes

From the distribution of velocities near the central regions of some galaxies, we can assume that a *supermassive black hole* (SMBH) is located at the center of all galaxies. Also, our galaxy hosts a SMBH at its very center (see, e.g., the review given by Kisselev et al. [125]). How can such a SMBH be detected? This follows from the extreme Keplerian motion near the center of Galaxies that can only be explained by a strong mass concentration. One of the first evidences from stellar dynamics of SMBH in our neighboring galaxy M31 was given by Dressler and Richstone [52], using observation of Ca II triplet (849.8, 854.2, and 866.2 nm), which comes mainly from the old stellar population of the nuclei of galaxies. They observed a full width at half maximum of these lines,[5] 0.22 nm, which corresponds to a velocity dispersion of 57 km/s. They estimated the central black hole mass in the nucleus of M31 at $6–7\times10^7\,M_\odot$.

Generally, SMBH correspond to masses of about $10^6\,M_\odot$. By inserting this mass in equation (3.5), we see that the radius of these SMBHs is still very small–less than 1 AU.

3.3.2.3 Active Galaxies

Supermassive black holes in the center of galaxies are thought to be the cause of extremely active types of galaxies. Active galaxies are unlikely to harbor habitable planetary systems due to strong high energetic radiation, gravitational perturbations, etc. Let us shortly describe them:

- *AGN, active galactic nuclei* galaxies.
- *Seyfert galaxies*: these are galaxies with extremely bright nuclei, their spectra have very bright emission lines of H, He, N, and O. The Doppler broadening of the emission lines implies Doppler velocities up to 4,000 km/s. These are believed to originate near an accretion disk surrounding the central black hole.
- *Quasars*: they show a high redshift, thus, they must be at large distances, but most of them look like point sources. Their brightness is powered by accretion of material into a SMBH in the nucleus. Thus, they are galaxies but with extremely active nuclei. The quasar 3C273 in the constellation of Virgo has absolute magnitude of 26.7. That means that at a hypothetical distance of 10 pc, this quasar would be of the same brightness as the Sun appears.

[5] This provides a measure for the line width.

Concerning habitability, it is important to answer the questions: why are some galaxies very active and some are almost nonactive? Could our Galaxy transform into an active galaxy? What happens during the motion of the solar system around the galactic center? The changing cosmic neighborhood because of this motion and other effects due to the passage of the solar system through the galactic plane will be discussed in Sect. 7.1

3.4 The Sun – Our Star

3.4.1 Overview

The Sun is the star we live with and it is the only star on which we can directly observe details on its surface. The importance of the Sun for life on Earth was already recognized by ancient cultures; in many cultures, the Sun is a god. From the astrophysical point of view, it is a prototype and important for comparison with other stars. Moreover, the Sun is the driver for weather and climate on our planet and without its luminosity the Earth would be a cold and icy world. We give the main properties of the Sun.

Distance	150×10^8 km
Mass	333, 000 Earth masses = 1.98×10^{30} kg
Radius	109 Earth Radii = 6.959×10^8 m
Luminosity	$3.826\times10^{26}\,\mathrm{W} = 3.826\times10^{33}\,\mathrm{erg/s}$
Apparent magnitude	$m_V = -26^{\mathrm{m}}.87$
Effective temperature	5,777 K
Rotation period	25.38 days

The Sun contains nearly 99.9% of the whole mass of the solar system. If the Earth would be placed at its center, the orbit of the moon would still lie within its body.

The solar constant S is the amount of solar energy that is received at the distance of the Earth on an area of 1 m^2 per s, $S = 1.36\,\mathrm{kW\,m^{-2}s^{-1}}$. However, this value can be only measured outside of the Earth's atmosphere.

The Sun is the driver of our weather and climate system. The energy is produced in the solar interior and transported outward. To understand how the solar luminosity is generated, the solar interior has to be examined. The internal structure of the Sun can be determined by theoretical models, helioseismology, or neutrino measurements. Helioseismology investigates the propagation of observed wave modes and derives from that the internal structure of the Sun (see, for example, the review given by Gizon and Birch [71]).

In the solar core, energy is produced mainly by the conversion of four protons (p) into a helium nucleus and two neutrinos (ν_e).

$$4p \rightarrow {}^4\mathrm{He} + 2e^+ + 2\nu_e \tag{3.9}$$

The mass of the resulting He is smaller than the mass of the four protons and, therefore, a mass Δm is converted into energy according the Einstein's equation:

$$E = \Delta m c^2 \tag{3.10}$$

The temperature at the *center* is about 1.4×10^7 K. The energy is produced in the form of photons (γ photons) that are absorbed, reemitted, and reabsorbed many times on their way throughout the solar interior to the surface. Due to this process, the radiation that is emitted presently was produced several 100,000 years ago in the solar interior. The energy is generated by thermonuclear fusion in the solar core and then transported outward first in the radiative zone. At about 200,000 km below the surface of the Sun, another energy transport mechanism sets in: *convection*. Hot gas rises, cools, sinks down, and the process starts again. The overshooting convection can be observed on the solar surface as a phenomenon called *granulation*. The surface of the Sun does not appear uniform but shows a cellular like pattern of a typical size of about 1,000 km. In the bright granules, matter rises, in the darker intergranular areas, it sinks down (Fig. 3.8, see, e.g., the review given by Muller [173]).

The extension of the three zones in the solar interior, core, radiative zone, and convection zone is roughly 1/3 of the solar radius $R_\odot$ each (Fig. 3.9).

The *photosphere* is the deepest layer we can observe but due to its small vertical extension of about 400 km, one has the impression that the solar limb appears sharp. Almost all radiation in the visible part of the spectrum is emitted from that layer. The mean surface temperature of the Sun is near 6,000 K and in the photosphere it falls down to about 4,000 K at the top (h = 400 km) and the density decreases from 10^{-7} to 10^{-8} g cm^{-3}. Thus, spectral lines are formed there in absorption; the deeper the line core, the higher it is formed. Most of the solar radiation in the visible is emitted from this layer, the maximum being about 500 nm.

The variation of several parameters in the solar interior is shown in Table 3.5.

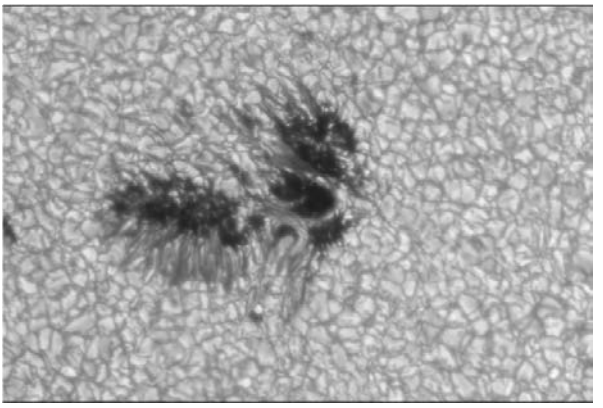

Fig. 3.8 Solar granulation with a sunspot. La Palma, Hanslmeier, Sobotka, 2003

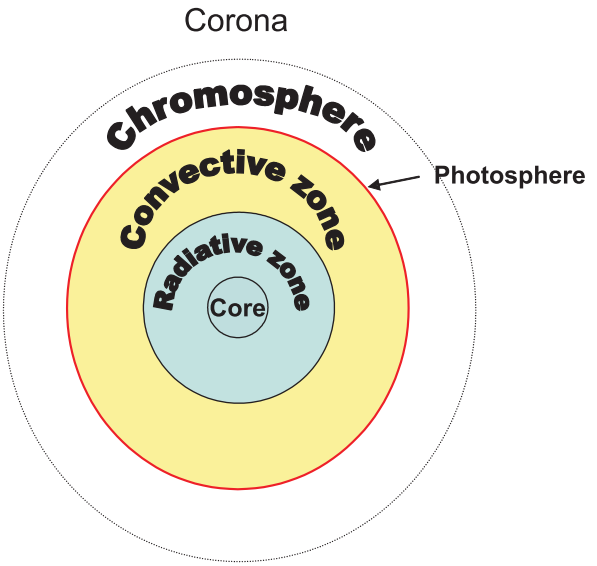

Fig. 3.9 The structure of the Sun: interior (core, radiative zone, and convective zone) and atmosphere (photosphere, chromosphere, and corona). The extension of the core, radiative zone, and convection zone is about 1/3 of the solar radius for each

Table 3.5 Solar model: variation of temperature, luminosity and fusion rate throughout the Sun

Radius fraction in $R_\odot$	Radius $[10^9]$ m	Temperature $[10^6]$ K	Luminosity %	Fusion rate[J/kg s]
0	0.00	15.7	0	0.0175
0.09	0.06	13.8	33	0.010
0.12	0.08	12.8	55	0.0068
0.14	0.10	11.3	79	0.0033
0.19	0.13	10.1	91	0.0016
0.22	0.15	9.0	97	0.0007
0.24	0.17	8.1	99	0.0003
0.29	0.20	7.1	100	0.00006
0.46	0.32	3.9	100	0
0.69	0.48	1.73	100	0
0.89	0.62	0.66	100	0

The *solar atmosphere* is a very interesting object to study because of several reasons. First, the temperature above the photosphere increases (see also Fig. 3.10): in the chromosphere (extension $\sim 10^4$ km); it increases to several 10^4 K; in the corona (extension several $R_\odot$), it increases to several million K. There are theories to explain this increase but, still, not all details are understood completely. Second, because high energetic phenomena take place in the chromosphere (flares) and corona (flares, CMEs (coronal mass ejections)) and magnetic fields determine the dynamics of matter as can be seen by loop structures on images.

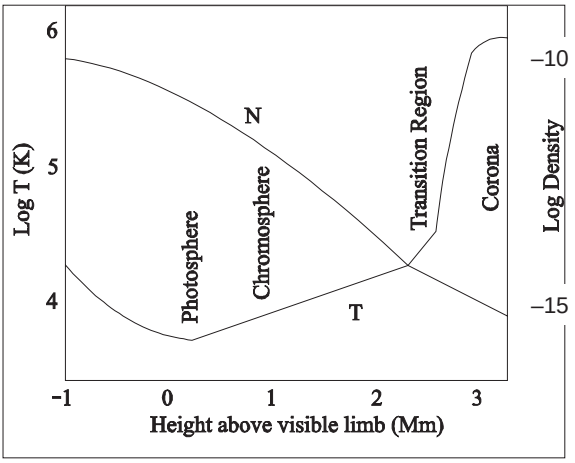

Fig. 3.10 The structure of the solar atmosphere and the variation of density and temperature. Note that the temperature increases from the solar surface (6,000 K) to several million K in the solar corona

In the corona, the density is extremely small, $10^{-18}\,\mathrm{g\,cm^{-3}}$. The structure of the corona is governed by magnetic fields. Plasma processes associated with solar activity take place on small temporal and spatial scales and reveal themselves by electron acceleration (up to a few MeV), which can emit radio radiation. These are called nonthermal electrons. Radio radiation from the Sun was first detected in 1942.

3.4.2 Solar Observations

The main problem for solar observations is its brightness (though when observing in very narrow bandwidths, there might be not enough light!). Big solar telescopes are often evacuated to minimize the heating of the telescope that would lead to a blurring of the image. In a vacuum tower telescope like the German Vacuum Tower Telescope (VTT) at the Observatorio del Teide, Tenerife, the solar light is caught by coelostat mirrors deflecting it into a vertical evacuated vacuum tank in which the further optics is place (Fig. 3.11). Smaller telescopes are used for patrol observations.

Since the Earth's atmosphere absorbs UV radiation (mainly by formation of ozone O_3) and X-rays (mainly by O_2), only telescopes from satellites (in former times also balloon flights) permit to observe the Sun in these wavelengths.

The solar activity is monitored from ground and space in order to understand the complex processes observed in various parts of the solar atmosphere. Currently, there are several space missions active like SOHO (**So**lar and **H**eliospheric

Fig. 3.11 The Observatorio del Teide in Spain, Tenerife. The coelostat mirrors have a diameter of 80 cm, the main mirror in the vertical evacuated tank 70 cm. The telescopes are operated by Germany (KIS, AIP, Göttingen) and Spain (IAC)

Observatory), TRACE (**T**ransition **R**egion **A**nd **C**oronal **E**xplorer, YOHKOH, HINODE, STEREO (**S**olar **Te**rrestrial **Re**lations **O**bservatory), ULYSESS, and RHESSI (**R**euven Ramaty **H**igh **E**nergy **S**olar **S**pectroscopic **I**mager).

Due to high temperatures in the corona and chromosphere, these layers radiate strongly in the UV and X-rays and can be observed from space-based telescopes (in former times these layers were only observable during total solar eclipses). The observation of eruptive processes in the chromosphere (prominences and flares) can be made also from ground-based observatories, e.g., in the light of the hydrogen line Hα.

The solar corona can also be studied at radio wavelengths. Because of the plasma frequency

$$\nu_0 = \sqrt{\frac{e^2 N_e}{\pi m}} = 9 \times 10^3 \sqrt{N_e} \tag{3.11}$$

and the decrease of the electron density N_e with increasing coronal height, one sees deeper into the corona (i.e., nearer to the solar surface) at higher frequencies.

Summarizing, we can state that a tomography of different layers of the solar atmosphere can be made by observing in different wavelengths: photosphere (visible radiation), chromosphere (Hα, UV), corona (EUV (extreme UV) and X-rays, radio emission). This is important to remember when considering stellar activity. Information on eruptive processes on stars is obtained by studying them at the respective wavelengths. Such eruptive processes strongly influence the planets near them.

3.4.3 The Variable Sun

The Sun determines our climate. The variability of a mother star, therefore, strongly influences on the habitability on planets due to its varying radiation (e.g., UV radiation, particle emission, magnetic field disturbances, etc.). A better knowledge of the variability of our host star enables us to understand stellar variability in general because only in the case of the Sun, fine structure can be observed directly.

The Sun is a variable star, but, compared with other variable stars, its variability is of extremely small amplitude. In 1610, for the first time, telescopic observations were made by Galilei and others and he detected sunspots, but the first recordings of sunspots date back to around 800 BC in China (big sunspots can be seen with the unaided eye). In 1843 Schwabe discovered the 11-year cycle of sunspot numbers.

3.4.3.1 Phenomena of the Active Sun

Associated with the sunspot number are other phenomena of solar activity like flares, prominences, CMEs, and the solar wind. We will briefly discuss these phenomena:

- Sunspots: they appear dark in the solar photosphere because the temperature in their central regions (umbrae) is about 2,000 degrees below the surrounding photosphere (about 6,000 K). If the Sun would consist of only one huge sunspot, then it would be still as bright as the full moon. The dark umbra is surrounded by a filamentary penumbra. The reason why sunspots are darker and, therefore, cooler are intense magnetic fields (Fig. 3.12). The magnetic field inhibits the convective transport of hot plasma to the surface. In the umbra, the field lines leave the solar surface nearly vertically becoming more inclined in the penumbra. Since there are no magnetic monopoles, many sunspots appear, therefore, as bipolar magnetic groups. The formation of a bipolar group occurs when magnetic ropes break through the photosphere. Large sunspot groups are visible over several solar rotations. The local magnetic field in the sunspots can be as strong as 4,000 Gauss,[6] the average solar magnetic field is around 1 Gauss. [7]
- Faculae: brighter than the surrounding photosphere, the temperature is higher. They are always associated with sunspots. By taking spectroheliograms at the wavelengths of Hα[8] or ionized Ca, one can observe faculae also in the chromosphere.
- Flares: energy that corresponds to 10^9 megatons of TNT or up to 6×10^{25} J[9] is released within a few minutes. They occur near active regions, i.e., sunspot groups, along the neutral line (that is the dividing line between areas of oppositely

[6] 1 Gauss = 10^{-4} Tesla.

[7] Average field strength on the surface of the Earth: 0.5 Gauss.

[8] This line is formed by the transition of the electron from level $n = 2$ to level $n = 3$ in the hydrogen atom.

[9] The Hiroshima atomic bomb had a power of 20 kt TNT.

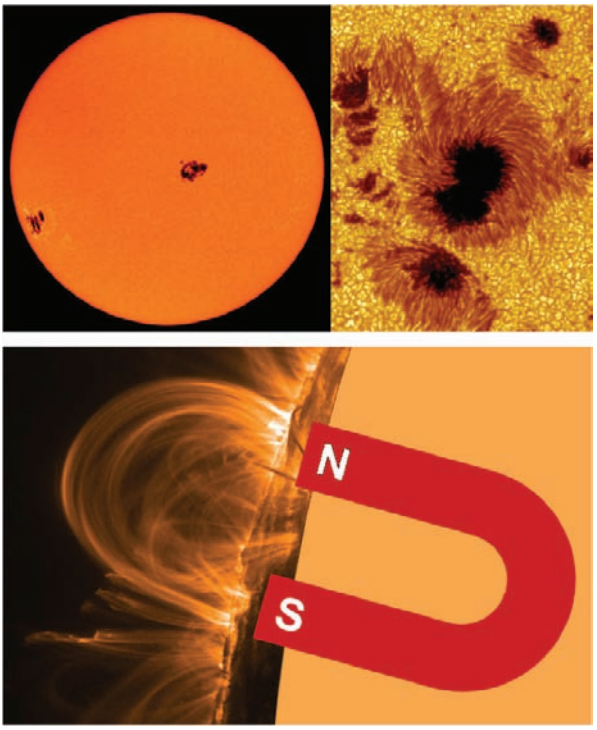

Fig. 3.12 Sunspots (*above*) and the extension of the magnetic field lines into the corona as seen by the distribution of matter (*below*). Photo Credits: ESA/NAS SOHO Mission, NASA TRACE Mission, Swedish La Palma Telescope

directed magnetic fields). During a flare event energy is released in form of high energy electromagnetic radiation (gamma-rays, X-rays, and EUV) and energetic particles (electrons and protons). The biggest flares are classified as X class; M flares have 1/10 of the energy of X flares. Flares can also be observed in the Hα spectral line.

- CMEs (Fig. 3.13): these are the coronal mass ejections–huge bubbles of gas that are ejected from the Sun in the course of several hours. They can only be detected by a coronograph in which the disk of the Sun is occulted artificially and the outer corona can be studied then (ideally from space where there is no stray light). The CMEs disrupt the flow of the solar wind and produce disturbances that also reach the Earth. They are often associated with solar flares and prominence eruption but they can also occur independently.
- Solar wind: a stream of charged particles in all directions, the velocities are about 400 km/s. There exists also a high stream component (800 km/s) that emerges from coronal holes and a low stream component (300 km/s) from the coronal streamers.

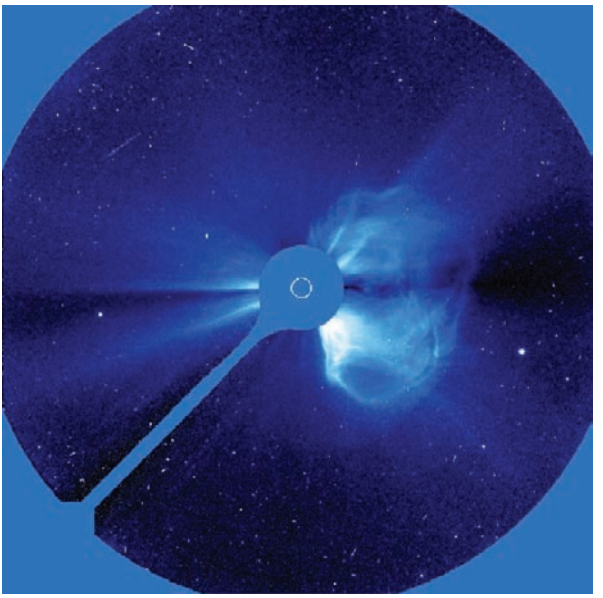

Fig. 3.13 A CME observed by the solar satellite SOHO. Note that the solar disk is occulted (its size is indicated by the inner bright ring). Photo Credit: SOHO ESA/NASA

- Prominences: dense clouds of material, their structure is formed by the magnetic field. They are suspended in complex, often loop-like forms over the surface by magnetic field.

All these above mentioned phenomena are variable and basically, the energy comes from a reconnection of magnetic field lines. In the photosphere, these field lines are frozen in, which means that they move with the motions caused by convection there and, thus, flux cancellation with energy release can happen in the higher corona.

The number of flares, CMEs, and sunspots varies nearly synchronously and, therefore, the solar activity cycle refers to all these phenomena.

3.4.3.2 Space Weather

The effects of these solar events on the Earth and the space near Earth[10] can be summarized as follows:

- Flares, CMEs → highly energetic particles in the solar wind → biochemical hazards to manned spaceflight.
- X-ray flares → radiation → increased ionization in the atmosphere, increase the drag on low orbiting satellites, interfere with radio communications, changes in the atmosphere.

More details can be found in Hanslmeier [91].

[10] This is often called as space weather.

3.4.4 Solar Activity Cycles

3.4.4.1 Long Term Periods

The activity of the Sun shows periodic cyclic behavior. The most prominent cycles are as follows:

- 11 year cycle, also called *Schwabe cycle*.[11]
- 22 year *Hale cycle*: the magnetic field of the Sun reverses during each Schwabe cycle. The magnetic cycle of the Sun is, thus, 22 years.
- 87 year: *Gleissberg cycle* (between 70 and 100 years), modulation of the Schwabe cycle.
- 201 years Suess cycle.
- 2,300 y: *Hallstatt cycle*.

As it is seen from Table 3.6, there were longer periods of enhanced solar activity as well as longer periods with no solar activity.

3.4.4.2 Solar Acitivity Proxies

Since observations of sunspots are available for only about 400 years, other *proxies of solar activity* have to be used to derive solar activity in the past. These include ^{14}C *isotope* measurements. This isotope is formed when more energetic particles penetrate through the magnetic field of the Earth. The extension of the solar magnetic field and the solar wind propagation over the whole planetary system defines the *Heliosphere*. The heliosphere provides a certain protection from energetic galactic particles, also called galactic cosmic rays (GCRs). Therefore, less ^{14}C is produced during phases of a strong heliospheric shield, i.e., when solar activity is maximum.

Table 3.6 Variations of Solar activity

Event	Start	End
Ooort Minimum	1040	1080
Medieval Maximum	1100	1250
Wolf Minimum	1280	1350
Spörer Minimum	1450	1550
Maunder Minimum	1645	1714
Dalton Minimum	1790	1820
Modern Maximum	1950	?

[11] From his observations made between 1826 and 1843, *Schwabe* found the solar cycle.

3.4.4.3 Solar Activity and Climate on Earth

The sensitivity of the climate on Earth to solar variations is higher for longer cycles. We have to take into account the thermal inertia of the ocean which damps high frequencies and there is a general lag of 2.2 years. It seems that the climate is 1.5 times as sensitive to the Hale cycle than to the Schwabe cycle. For example, Raspopov et al. [201], studied conifer tree rings for the period 1458-1975 and found climatic oscillations with periods of 20–25 years, which they attributed to a solar forcing.

The longer term variation of solar activity is given in Fig. 3.14.

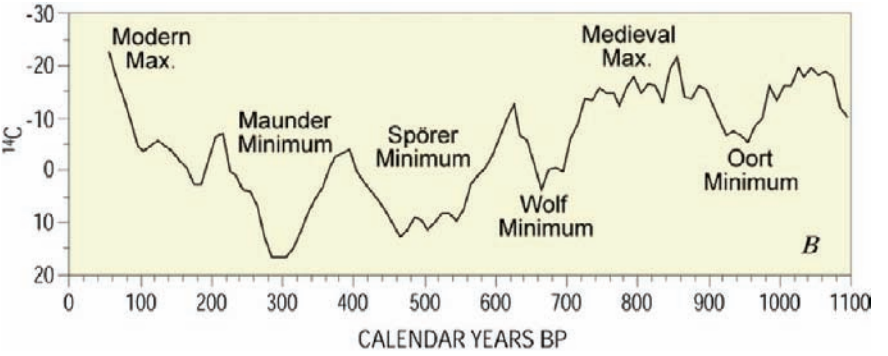

Fig. 3.14 Solar activity derived from the ^{14}C concentration as a proxy

The connection between periods of high or low solar activity and climate on Earth seems to be in such a way that prolonged periods of high solar activity coincide with warmer climate on Earth, whereas periods of prolonged low activity coincide with a lower global temperature on Earth. Today, it is well known that the *Maunder Minimum* (1645–1715, during that interval no sunspots were recorded) coincided with the coldest part of the so-called *little Ice Age*. During the Maunder Minimum, in Europe and North America, the winters were extremely cold and the summers wet and this had negative effects on the crops. The Thames river and the canals and rivers of the Netherlands often froze; glaciers in the Alps advanced.

The concrete effects on habitability of large solar variations will be discussed later.

Chapter 4
Planetary Systems

The year 1995 has seen a very important discovery. For the first time, a planet was detected that does not belong to our solar system. Since that time, the methods to find extrasolar planets have been refined and up to the year 2008, more than 250 extrasolar planets were listed. Since planets only reflect light from their central star, they are extremely difficult to detect:

- The radiation from the central star is orders of magnitudes higher than the radiation of the orbiting planet.
- As seen from Earth, the planet is generally quite close to the central star.
- These observational difficulties cause strong bias since it is more likely that massive planets will be detected than small faint ones.

The discovery of the first extrasolar planet can be compared with the Copernican Revolution that not the Earth but the Sun is the center of our planetary system and the discovery of Hubble, that the universe is expanding. Since 1995 it is evident that planetary systems are common in the universe and this certainly has a big impact on the question of whether there are other habitable planets or not.

We start this chapter with a brief overview about our planetary system, the solar system.

4.1 The Solar System: Overview and Formation

4.1.1 Planets — How Are They Defined?

What is the difference between a star and a planet? A first answer could be that planets do not shine by themselves, they reflect light from their host star. A more precise definition includes the mass of the objects.

- Bodies with masses $> 0.075\,M_\odot$ reach core temperatures $\sim 10^7$ K where hydrogen fusion sets in. Note that a mass $> 0.075\,M_\odot$ equals to about 75 times the mass of Jupiter (also written as 75 M_J).
  Mass of Jupiter: 1 M_J= 318 M_E= 2×10^{27} kg, M_E is the mass of the Earth.
  Stars at the lower mass edge are expected to have a surface temperature of only 2,000 K.

A. Hanslmeier, *Habitability and Cosmic Catastrophes*, Advances in Astrobiology and Biogeophysics, DOI 10.1007/978-3-540-76945-3_4,

- Brown Dwarfs: their mass range is between $13\,M_J$ and $75\,M_J$. The distinction between the brown dwarfs and the even less massive planets is that the core temperatures of brown dwarfs can reach 10^6 K and higher. Thus, short intervals of nuclear burning of deuterium and lithium are likely. Young brown dwarfs are similar to very low massive stars. Older brown dwarfs are similar to a giant Jupiter except that their surface temperature is much higher (900 K[1]). Their atmosphere contains large amounts of water vapor and methane. The size of brown dwarfs is similar to that of Jupiter. From the dynamical point of view it has to be stressed that brown dwarfs are predominantly free-floating objects like stars. They can be found by IR surveys and more than 100 objects are known.
- Planets: objects with masses less than $13\,M_J$ are classified as planets.

The lower mass limit for planets is such that he has "cleaned" his orbit – there are no small particles left – and that he has spherical shape.

A planet is a spherical object with mass less than $13M_J$.

The astronomical nomenclature for companion stars are the capital letters A, B,...; for planets, the letters a,b,... are used.

4.1.2 Overview and Formation of the Solar System

4.1.2.1 Objects in the Solar System

The main body of the solar system is the *Sun*. The Sun contains 99.8% of the system's known mass. There are two big planets, *Jupiter* and *Saturn*, that contain more than 90% of the remaining mass. The eight *great planets* revolve around the Sun in the same sense and almost in the same plane — the plane of the Earth's orbit— which is called the ecliptic. Thus, all planetary orbits lie near the ecliptic, which was already known by ancient cultures. Other members of the solar system are the *dwarf planets* (e.g., Pluto), *comets*, *asteroids*, and *interplanetary matter*.

Distances in the solar system are usually given in *astronomical units* (AU).

1 AU is the mean Earth-Sun distance $= 149.6\times10^9$m.

Except for Venus and Mercury, all planets have natural satellites.

The whole solar system is enveloped by a cloud of cometary like bodies, the *Oort cloud*. This spherical cloud extends about $50,000$ AU to $100,000$ AU from the Sun.[2] The Oort cloud is supposed to be a remnant of the primordial solar

[1] The surface temperature of Jupiter is 130 K.

[2] This corresponds to 1/4 of the distance to the nearest star Proxima Centauri.

nebula from which the Sun originated 4.6 billion years ago by a collapsing process. The question why this Oort cloud is spherical and not concentrated near the ecliptic can be answered as follows: it is assumed that the objects of the Oort cloud formed originally much closer to the Sun but because of gravitational interaction with the young gas planets, they were ejected into elliptical or parabolic orbits and, thus, scattered out of the ecliptic plane. Thus, the formation and presence of large gas planets in an early stage of planetary system formation in general is essential to get rid of the many small bodies that would be hazardous in the inner planetary system to bombard the planets there and to extinct life.

4.1.2.2 Presolar Nebula

The presolar nebula had a diameter between 7,000 and 20,000 AU. An example of a protoplanetary disk is given in Fig. 4.1. This protoplanetary disk is observed in a star-forming region in the *Orion nebula* at a distance of about 1,600 Lyr. As it collapsed, it started to rotate faster because of angular momentum conservation. Most mass accumulated at the center. Due to rotation, the nebula flattened to a protoplanetary disk of about 200 AU and in its central region, a protostar has evolved. A similar scenario led to the formation of the protosun. T Tauri stars are believed

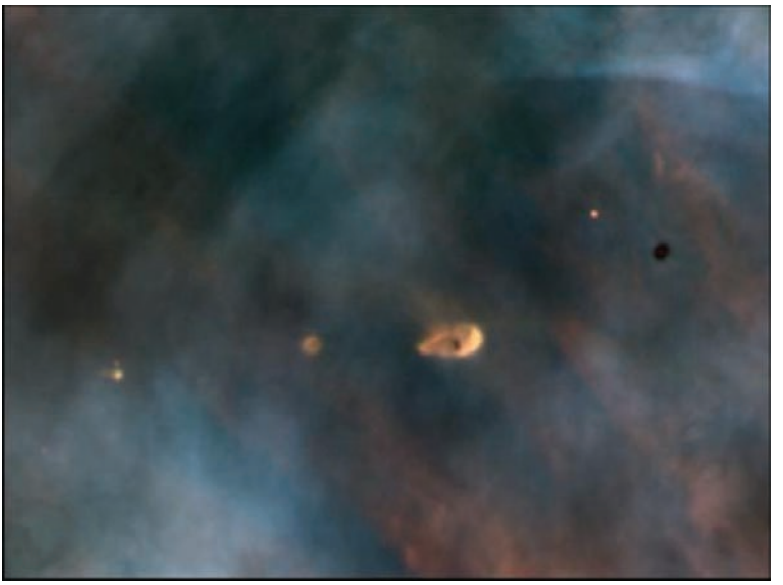

Fig. 4.1 An example of a protoplanetary disk observed in the Orion nebula M42. The Orion nebula is at a distance of 1,600 Lyr and the size of the field of view in the image is 0.14 Lyr which corresponds to about 9,000 AU. Photo credit: HST

to be in the same stage as the protosun. They are hot and luminous and their luminosity is generated by contraction. This phase is also called the T Tauri phase of the Sun and it lasted about 100 million years. At the end of the T Tauri phase, pressure density and temperature in the center were high enough that nuclear fusion has started. Since that time, the Sun produces its energy by nuclear fusion. The remaining cloud of dust served to form the planets that started as dust grains which formed planetesimals (up to 5 km in size) and due to frequent collisions, finally, the planets evolved.

It is also evident that in the inner solar system, because of the higher temperature, volatile molecules like water or methane could not condense. Only bodies with high melting points (e.g., silicates and metals) were formed there — the terrestrial planets Mercury, Venus, Earth, and Mars. Farther away from the Sun, beyond the so-called *frost line*, volatile compounds remained solid and Jupiter and Saturn became gas giants. Uranus and Neptune gathered less material. We assume that their cores are made mostly of ices (hydrogen compounds).

The starting point of the Sun as a main sequence star was also important for the cleaning of the protoplanetary disk. Because of strong solar wind (by a comparison with young solar – like stars, we know that the solar wind in the early phases of solar evolution had been by a factor of 100 stronger than today), gas and dust in the protoplanetary disk were blown into interstellar space. The accretion of planets therefore stopped.

4.1.3 Instable Interstellar Clouds

4.1.3.1 What Triggers the Collapse of a Cloud?

The formation of the solar system can be described briefly by a collapse of an interstellar cloud consisting of dust and gas. Such clouds must be triggered by an external agent in order to collapse. External agents could be shock waves due to a nearby supernova explosion (Fig. 4.2 and Fig. 4.3), disturbances caused by the rotation of the cloud about the center of the galaxy, density waves (that are at least partly responsible for the spiral structure of our galaxy), interaction of gas clouds in colliding galaxies (star burst galaxies), and UV shock wave from OB stars in H II regions.[3]

The processes are complicated. For example, if a very luminous star in an interstellar cloud has formed, it could either trigger the formation of further stars or inhibit the formation of protostars by eroding away the surrounding material.

A cloud starts to *collapse* when the gravitational attraction becomes larger than the internal gas pressure, this is also called *Jeans criterion* and can be derived from the condition that

$$E_{\text{grav}} + E_{\text{kin}} < 0 \tag{4.1}$$

[3] These are regions where the temperature is high enough that hydrogen is ionized.

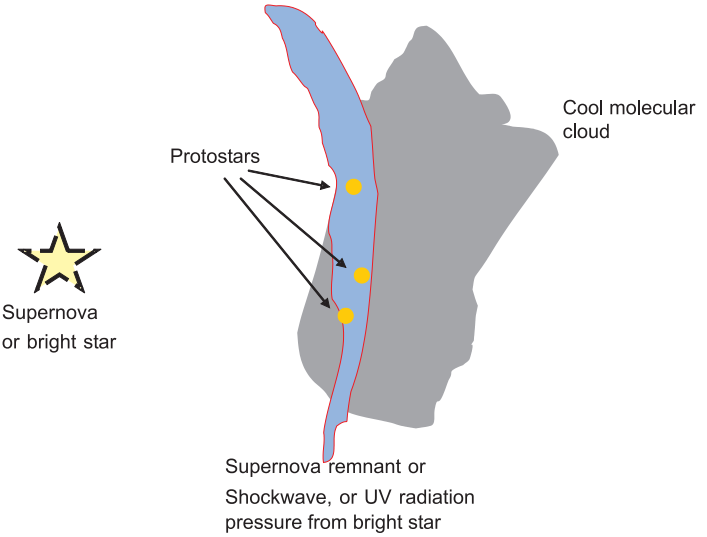

Fig. 4.2 Star formation induced by a shock wave (either from a supernova explosion or by radiation pressure of a bright star) entering a cool molecular cloud

• For a spherical cloud with mass M and radius R the gravitational energy is

$$E_{\text{grav}} = \int_0^M \frac{GM}{r} dm = -\alpha \frac{GM^2}{R} \tag{4.2}$$

where α depends on the degree of the central condensation or internal density distribution: if r denotes the distance from the center of the cloud, then $\alpha = f(\rho(r))$. The more the density increases toward $r \to 0$, the greater α. For uniform density, $\alpha = 3/5$. One often uses $\alpha = 1$.

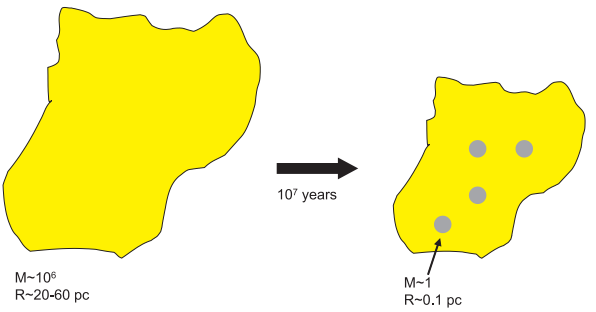

Fig. 4.3 A molecular cloud collapses in free fall, its temperature remains nearly constant and fragmentation sets in, and dark globules ("Bok Globules") are visible against the bright nebula. The process to form these globules lasts about 10^7 yr

• Now, let us estimate the kinetic energy: if T denotes the average cloud temperature and μm_H the average mass of a cloud particle, then

$$E_{\text{kin}} = \frac{3}{2}\frac{M}{\mu m_H} kT \tag{4.3}$$

Thus, for a collapse of a cloud with radius R to occur, there are two possibilities: *Jeans mass*:

$$M_{\text{cloud}} > M_J = \frac{3kT}{2G\mu m_H} R \tag{4.4}$$

Jeans density:

$$\rho_{\text{cloud}} > \rho_J = \frac{3}{4\pi M^2}\left[\frac{3kT}{2G\mu m_H}\right]^3 \tag{4.5}$$

The typical parameters for the interstellar matter are as follows. Particle numbers $\sim 5 \times 10^6\,\text{m}^{-3}$. The densities of the molecular clouds where stars form are much higher: $n \sim 5 \times 10^9\,\text{m}^{-3}$, typical dimensions are 10 pc, temperature about 10 K.

4.1.3.2 Stages of Star Formation

1. Contraction: as soon as the Jeans's criterion is fulfilled, the cloud starts to contract → release of gravitational energy E_{grav}. This released energy increases the temperature of the cloud inside. Due to the temperature increase, the molecules there (mainly molecular hydrogen H_2) are excited into rotational states. The de-excitation emits photons at low energies → radiation in the IR and mm-wavelength domain. To these wavelengths, the cloud is transparent, the photons can escape → cloud is cooled and contracts further.
2. Fragmentation: subregions of the clouds reach Jeans density and start to contract independently. This means that stars always form in clusters — most of these clusters dissolve quickly. It is not clear where the stars are located now that formed out of the same interstellar cloud as our Sun.
3. Formation of protostars. Consider a protostar with 1 solar mass, radius $10^{15}\,\text{m}$, and with a Jeans density $10^{-16}\,\text{kg}\,\text{m}^{-3}$. The radius after gravitational contraction will then be $R_2 \sim 10^{11}\,\text{m}$ and the timescale for the contraction is $\tau \sim 20,000\,\text{yr}$ (Fig. 4.4).
4. Hydrostatic equilibrium: after all hydrogen is ionized, the internal pressure rises, which slows down contraction and finally leads to hydrostatic equilibrium. In this phase, stars spend most part of their lifetime.
5. Lower limit of mass for nuclear reactions to occur: $M \sim 0.08\,\text{M}_\odot$ objects with less masses do not become stars because nuclear reactions cannot start.
6. Formation of planetary systems: due to rotation of the disk around the protostar (in the same sense as does the protostar), the disk flattens. Often, a bipolar outflow is observed. In the accretion disk, the density increases and the formation of particles growing in size sets in. The formation of an accretion disk seems a critical point because a double star could also be formed out of

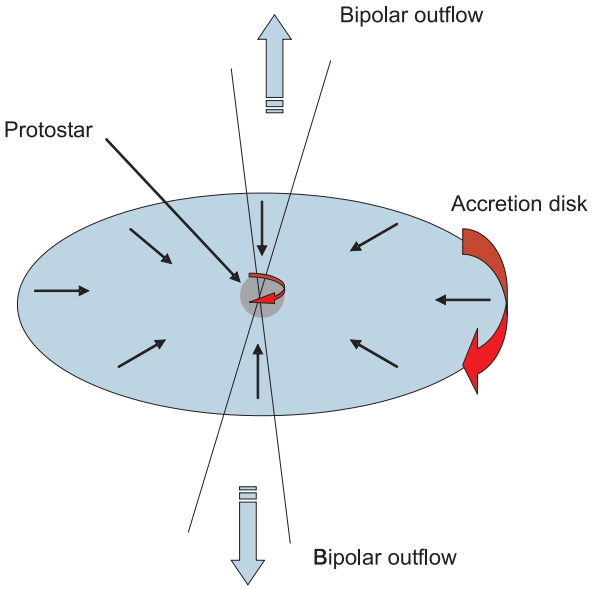

Fig. 4.4 An accretion disk has formed that rotates in the same sense as the protostar. In this flattened disk, planets are formed. Matter is expelled by a bipolar outflow

it. The dust particles in the disk collide with each other and form larger particles. Boulders are formed (small asteroids). Then, a run-away growth starts. The larger ones start to collect the smaller particles. Calculations show that in the inner solar system, objects between the size of asteroids and up to the size of the Moon could have been formed this way and in the outer solar system, up to the size of 15 times the Earth's size. This accretion of the planetesimals took $10^5 - 20 \times 10^6$ yr. The larger the distance from the sun, the longer this process lasted.

7. Stellar wind: about 1 million years after the nebula cooled, the star generates a very strong stellar wind. This would sweep away all of the gas left in the protoplanetary nebula. If a protoplanet was large enough, its gravity would have been able to attract the nebular gas and it would become a gas planet. If not, the planet would remain rocky or icy.
8. After $10^7 - 10^8$ yr, the planets of the solar system have been formed. The time of the *heavy bombardment* started and their surface became heavily bombarded by the remaining small boulders.

Summarizing the process of star formation — it starts with a free fall of matter, homologous contraction of a nebula (Jeans criterion):

$$\tau_{\text{freefall}} = \left(\frac{3\pi}{32 G \rho} \right)^{1/2} \tag{4.6}$$

$\rightarrow \tau_{\text{freefall}} \sim 10^4\,\text{yr}$.

Then, the *Kelvin-Helmholtz* time scale becomes important during which the star becomes heated up to the point when nuclear reactions set in:

$$\tau_{\text{KH}} = \frac{3GM^2}{10R}L \tag{4.7}$$

$\rightarrow \tau_{\text{KH}} \sim 10^7\,\text{yr}$.

The main sequence lifetime of a star depends on its mass:

$$\tau_{\text{Nuc}} = \frac{0.1\frac{\Delta m}{m}Mc^2}{L} \tag{4.8}$$

We have assumed here that 10% of the stellar mass are converted to helium.

This model of star and planetary formation can explain characteristic facts of our solar system:

1. Most angular momentum is in the planets.
2. The orbits of the planets lie nearly in the same plane.
3. The planets orbit in the same direction.
4. Planets (Venus is an exception) rotate in the same direction (i.e., direction of revolution).
5. The rotation axis is nearly perpendicular to the ecliptic.
6. The distances of the planets can be expressed by en empirical law (Titius-Bode law).

4.2 Main Objects in the Solar System

4.2.1 Planets

4.2.1.1 Terrestrial Planets

The terrestrial planets Mercury, Venus, Earth, and Mars have nearly the same structure: a central metallic core (Fe, Ni) surrounded by a silicate mantle. On the surface, we find craters, mountains, and volcanoes. They possess secondary atmospheres (very thin in the case of Mercury). Secondary atmospheres are a result of internal volcanism outgassing and maybe of cometary impacts. The gas giants have primary atmospheres. These atmospheres are the result of the formation of the planet from the original solar nebula and, thus, reflect the composition of the nebula. Terrestrial planets are also called Earth-like planets.

Mercury is the closest planet to the Sun and the temperature on the dayside is up to 450°C and on the nightside upto –180°C. This extremely high temperature contrast is a consequence of lacking an atmosphere. *Venus* is sometimes referred to as “sister Earth” because it is similar in size. However, the conditions for life are not very promising. The surface pressure of the atmosphere is 90 times that of Earth’s[4]

[4] That corresponds to the pressure at a water depth of 900 m on Earth.

and, because of the dense atmosphere that consists mainly of the greenhouse gas CO_2, the surface temperature is about 460°C. Venus rotates in 243 Earth days in the direction opposite to its orbital motion and one revolution about the Sun lasts 224.7 days.

The *Earth* is the largest of the terrestrial planets. The mean radius is 6,371 km, the mass is 5.97×10^{24} kg, and the density is $5.5\,\mathrm{g\,cm^{-3}}$. The oceans comprise 2/3 of its surface. Over the whole evolution, the mass of land has increased steadily; during the past two billion years, the total size of the continents have doubled. Like the other terrestrial planets, the interior of the Earth can be divided into layers: the lithosphere is found at a depth between 0 and 60 km, the crust between 5 and 70 km, the mantle between 35 and 2,890 km, the outer core between 2,890 and 5,100 km, and the inner core between 5,100 and 6,371 km (density $\sim 12\,\mathrm{g\,cm^{-3}}$). The internal heat is produced by radioactive decay of ^{40}K, ^{238}U, and ^{232}Th. The rigid crust is very thin compared with the other layers; beneath oceans, it is only 5 km, beneath continents, $\sim$ 30 km, and beneath mountain ranges (Alps), it is up to 100 km. The atmosphere is composed of 78% nitrogen and 21% oxygen.

Mars is the second closest in similarity to Earth. On a nice summer day, the temperature on Mars may rise up to 0°C; however, during a Martian night, it may reach −100°C. There are channels on its surface, which are a hint that this planet underwent large climatic variations in the past with episodes of liquid water on its surface. Today, because of the low atmospheric pressure (only 1% of the pressure on the surface of the Earth), water cannot exist in liquid state on Mars.

4.2.1.2 Gas Giants

The largest planet in the solar system is *Jupiter* and like the other gaseous planets, it is composed primarily of H and He with no solid surface. Its diameter is 11 times that of the Earth. The temperature on Jupiter's cloud tops is -148° C . *Saturn* is known for its bright ring system (in fact all gaseous planets have ring systems). The rings are only several 100 m thick and consist of billions of ice particles most of them cm sized. *Uranus* has a tipped rotation axis, i.e., it seems to roll around the Sun rather than spinning. *Neptune* has an internal heat source as Jupiter and Saturn do. That means these planets gives off more radiation than they receive from the Sun.

Some important data of the planets are given in Table 4.1. In this table, D denotes the distance from the Sun, P the period of revolution about the Sun, R the radius of the planet, and P_{Rot} the rotation period.

4.2.2 Dwarf Planets

In August 2006, the International Astronomical Union (IAU) defined a new class of objects in the solar system, the dwarf planets. These objects must fulfill the four conditions:

1. Orbit around the Sun
2. Sufficient mass, which implies a hydrostatic equilibrium and a near spherical shape
3. Has not cleared the neighborhood around its orbit
4. Is not a satellite.

Pluto is one famous member of this class. Other objects are, e.g., *Ceres* (which was formerly classified as asteroid) and *Eris*. The diameter of Pluto and Eris is nearly the same—about 2,400 km; the semi-major axis of Pluto is 39.5 AU (high eccentric, orbital radius between 29.66 and 49.3 AU) and that of Eris 67.67 AU (high eccentric, orbital radius between 37.77 and 97.56 AU. Ceres has a diameter of 975×909 km and the semi-major axis is 2.76 AU. Pluto is also defined as a prototype of *Trans Neptunian Objects (TNOs)*.

4.2.3 Comets

The Greek word *kome* denotes hair of the head and the designation *kometes* means stars with hair. Often through history people believed that the appearance of comets means something dangerous like war or plagues.

After *Newton* established his law of gravity in 1705, *Halley* detected that the comets of the years 1531, 1607, and 1682 have very similar orbits and belong to the same object. This comet was named after Halley. It was a big triumph of celestial mechanics when the predicted reappearance of Halley's comet in 1758–9 could be observed. Its next appearance will be in 2061. The official designation is *1P/Halley* where the *P* stands for periodic comets.

4.2.3.1 Different Parts of a Comet

Nucleus: the size of the nucleus is in the order of several 10 km. It is composed mainly of rock, dust, and ice. Therefore, in 1960[5], *Whipple* (1906–2004) introduced

Table 4.1 Some important parameters of the planets in the solar system

Planet	D $\times 10^6$ km	P	R (km)	P_{Rot}
Mercury	57.91	88.0 d	4879 (0.38)	58.65*d*
Venus	108.21	224.7d	12742 (1.0)	–243.02 d
Earth	149.6	365.25 d	12742	23h56m
Mars	227.92	687.0 d	6780 (0.53)	24h37m
Jupiter	778.57	11.75 y	139822 (10.97)	9 h 55m
Saturn	1433.53	29.5 y	116464 (9.14)	10h 40m
Uranus	2872.46	84 y y	50724 (3.98)	–17.24 h
Neptune	4495.06	165 y	49248 (3.87)	16.11*h*

[5] Reference [267].

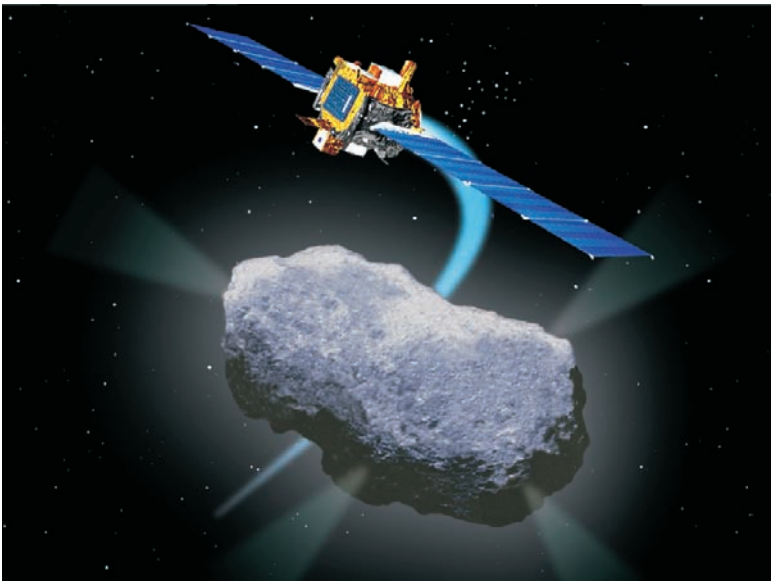

Fig. 4.5 Artist's concept of Deep Space 1 encounter with comet Borrelly. NASA/JPL

the designation *dirty snowball*. The ESA-Giotto mission (1986) found that cometary nuclei are extremely dark objects; comet Halley's nucleus reflects approximately 4% of the light, and the Deep Space mission 1 discovered that Comet *Borrelly*'s surface (Fig. 4.5) reflects only 2.4% to 3%.[6] The reason for such a low albedo is a cover of the surface by dark complex organic compounds. Solar heating drives off volatile compounds leaving behind heavy long-chain organics that tend to be very dark, like tar or crude oil. The very darkness of cometary surfaces allows them to absorb the heat necessary to drive their outgassing.

Coma: when the comet approaches the Sun (i.e., inside the distance of Mars), evaporation at the surface on the nucleus starts and a coma (cometary atmosphere) is formed. The extension of the coma is several 10^3 km.

Tail: besides the coma, very often a tail is formed that can extend up to 10^8 km. The tails consist of two components both always pointing away from the Sun. The particles forming a dust tail move around the Sun according to Kepler's laws and particles farther away from the Sun move slower.[7] Therefore, long dust tails appear curved. The light pressure from the Sun (photons have a momentum of $h\nu/c$) pushes the particles in the direction opposite to the Sun. The ion tail is formed by the ionized gas component and is strongly affected by the solar wind and it follows magnetic field lines rather than an orbital trajectory. The ion tail often has a bluish appearance. The dust tail is brighter than the ion tail. Comets also emit X-rays: when highly charged ions of the solar wind particles fly through a cometary atmosphere, they

[6] For comparison: asphalte reflects 7%.

[7] Like planets: i.e., Mars moves slower than the Earth.

collide with cometary atoms and molecules. In these collisions, the ions will capture one or more electrons leading to emission of X-rays and far ultraviolet photons.

Due to mutual perturbations of objects in the Oort cloud (maybe also due to perturbations of nearby stars), some of these objects move inside to reach the inner solar system and can then be observed as comets. On their journey to the inner solar system, they can be perturbed — mainly by the strong gravitational attraction of the two largest planets of the solar system, Jupiter (more than 300 Earth masses) and Saturn (about 95 Earth masses). Therefore, their orbit changes and they become short period comets. Some comets are moved into sungrazing orbits and they have been detected with the SOHO solar satellite — in August 2005, SOHO discovered its 1,000th comet. Others may be thrown out of the solar system forever.

A histogram of the aphelia of 2,500 comets is shown in Fig. 4.6. The comet families are a consequence of the perturbations of the giant planets.

4.2.3.2 Classification

A classification of comets may be made according to their orbital periods. *Short period* comets (also called periodic comets) reappear within less than 200 years, main-belt comets orbit within the asteroid belt. *Long period* comets may have

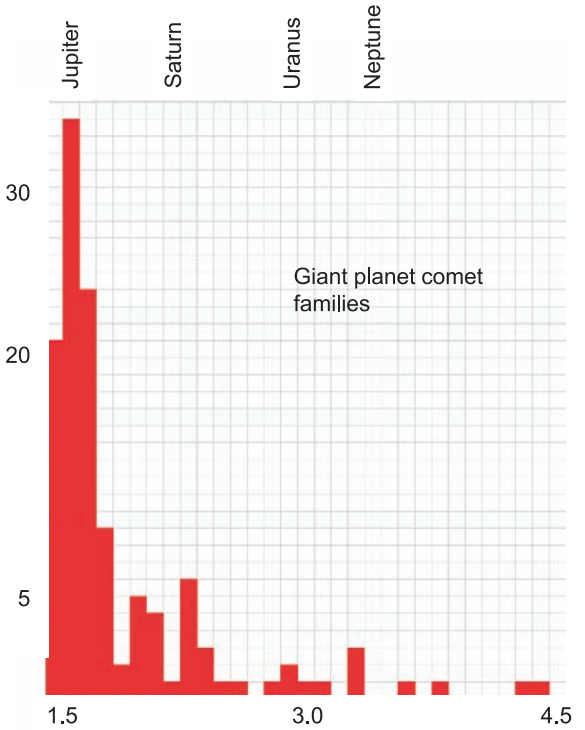

Fig. 4.6 Histogram of the aphelia of the 2005 comets, showing the giant planet comet families. The abscissa is the natural logarithm of the aphelion expressed in AU

.

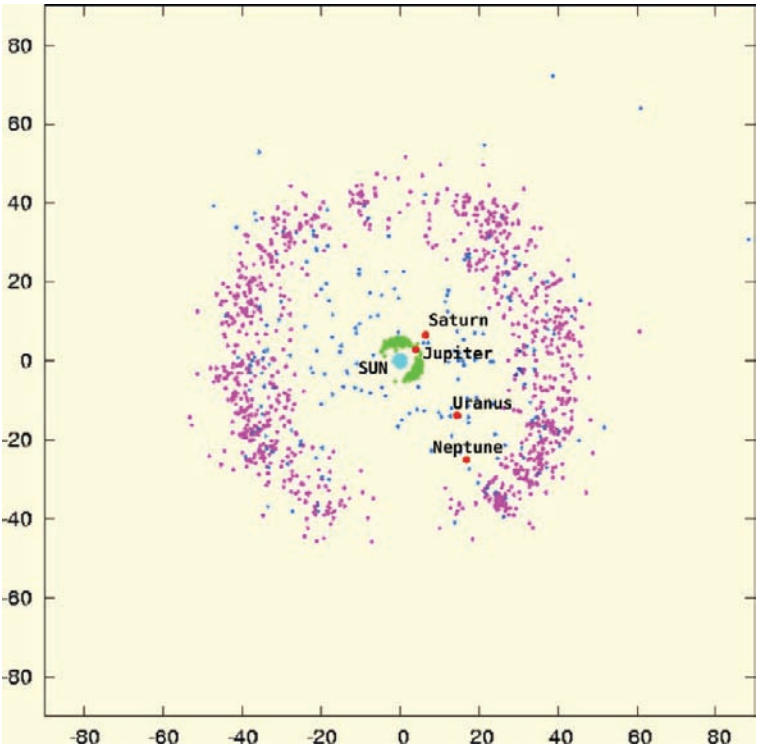

Fig. 4.7 The known objects in the Kuiper belt. Note the Trojan asteroids that form an equilateral triangle with Jupiter and the Sun. Courtesy: Wikimedia

periods of several 10^4 years but they still belong to the solar system. Due to perturbations (mainly by Jupiter), comets may get on a hyperbolic orbit, i.e., they will exit the solar system after perihelion passage.[8]

Short period orbit comets originate in the *Kuiper belt* (Kuiper, 1951). The Kuiper belt lies beyond the orbit of Neptune and extends from about 30–100 AU (Fig. 4.7). The first object that belongs to this belt was found in 1992; up to now more than 1,000 objects have been found, denoted as Kuiper belt objects (KBOs). *Pluto* and, e.g., *Quaoar* (discovered in 2002) are large KBO and the Neptunian satellite *Triton* is a captured KBO.

Long period orbit comets originate in the *Oort cloud*.

All objects whose distance from the Sun on an average is greater than Neptune are called *Trans Neptunian Objects* (TNOs). It is believed that there exist three categories of these objects:

- Kuiper belt objects, KBOs,
- Scattered objects, SDOs, and
- Oort cloud objects, OCOs.

[8] At its perihelion, a comet is nearest to the Sun.

Table 4.2 Large Kuiper Belt Ojects, KBOs

Object	Aph. distance (AU)	Perih. distance (AU)	Size (km)
(136472) 2005 FY9	53.0	38.5	1600–2000
(136108) 2003 EL61	51.5	35.1	~ 1500
Quaoar	44.9	41.9	1300
Eris	37.8	97.5	1200
Sedna	76	900 (!)	1700

It is interesting to note that a high fraction of TNOs are binaries with roughly equal masses and high eccentric orbits. Their separation is in the order of 100 R (radius of the primary). Funato et al. [67] discuss the formation mechanisms for such objects.

The *Cubewano objects* are KBOs with orbits beyond Neptune.

The *Centaurs* orbit the Sun between the orbits of Jupiter and Neptune. Saturn's moon Phoebe may be a captured Centaur.

Some large KBOs are given in Table 4.2.

A well – known example of a short period comet is *Encke* (2P/Encke, detected as the second known periodic comet by Encke in 1819). The aphelion distance of 2P/Encke is 4.11 AU and perihelion distance 0.33 AU; the last perihelion passage was in April 2007. The orbital period is 3.3 years.

The aphelia of short period orbit comets often coincide with a giant planet's semi-major axis forming a *family of comets*.

4.2.3.3 Designation and Orbital Parameters

The designation of comets is complicated. Until 1994, the designation was like, e.g., Comet 1969i, which was the ninth comet discovered in 1969. Once the comet had been observed through perihelion and its orbit had been established, the comet was given a permanent designation: year of its perihelion followed by a Roman numeral indicating its order of perihelion passage: i.e., 19969i $\rightarrow$ Comet 1970 II (the second comet to pass perihelion in 1970).

In 1994, the IAU approved a new system. Since that year comets are designated by the year, of their discovery followed by a letter indicating the half month of the discovery and a number indicating the order of discovery (the same system is used for asteroids). For example, 2007 D5 is the fifth comet discovered in the second half of February 2007. The prefix P/ indicates a period comet, C/ a nonperiodic comet, and X/ a comet for which no reliable orbit could be calculated. D/ indicates a comet which has broken up or been lost and A/ an object that was identified as a comet but is a minor planet.

As far as orbital parameters of comets are concerned, very often the following set of parameters can be found in the literature: q... perihelion distance in AU, e... eccentricity, i... inclination, P... orbital period in years, and T_J... Tisserand invariant.

The *Tisserand invariant* (Tisserand criterion) follows from the *restricted three body problem*. Let us consider three masses: m_1 = Sun, m_2 = Jupiter, and m_3 = comet. The comet does not influence the motion of the Sun and Jupiter, and moreover, $m_1 >> m_2$. Because of the conservation of the *Jacobi integral* in the restricted three body problem, the following criterion can be found: let $a, e,$ and i be the semi-major axis, eccentricity, and inclination of the cometary orbit before a close encounter with Jupiter and a', e', i' be these values after a close encounter. Then,

$$\frac{1}{2a} + \sqrt{a(1-e^2)}\cos i = \frac{1}{2a'} + \sqrt{a'(1-e'^2)}\cos i'. \tag{4.9}$$

How can we apply this criterion? whenever a new comet is discovered, the Tisserand parameter can be calculated:

$$T_J = \frac{1}{a} + 2\sqrt{a(1-e^2)}\cos i. \tag{4.10}$$

If this parameter has the same value as that of a previously observed comet, then it is quite likely that the new comet is, in fact, the same comet, but that its orbital parameters have changed since it was last observed due to a close encounter with Jupiter. In our case, the system of units is such that the semi-major axis of Jupiter is unity.

4.2.4 Asteroids

4.2.4.1 Discovery and General Properties

Asteroids, also called minor planets or planetoids, are objects smaller than planets and dwarf planets that orbit the Sun, most of them in the asteroid belt between Mars and Jupiter. In the General IAU Assembly in 2006, these objects together with the comets discussed above, are also designated as *Small Solar System Bodies (SSSBs)*. At that assembly, also, a new class of objects was defined: the dwarf planets to which Pluto belongs.

The first asteroid was discovered by *Piazzi* in 1801 and named *Ceres*. However, since 2006, this object is classified as a dwarf planet. Nevertheless, it was the first object detected between the orbits of Mars and Jupiter and within a short time, many objects smaller than Ceres were found. 2 *Pallas*, 3 *Juno*, and 4 *Vesta* were discovered over the next few years, with Vesta found in 1807.

Asteroids are an interesting class of objects since it is assumed that they are remnants of the protoplanetary disk, planetesimals, that have not undergone transformation since the formation of the solar system. Because of the large perturbation of Jupiter, the objects in the belt between Mars and Jupiter did not form a planet. It was a big surprise, however, to find that several asteroids have satellites (Fig. 4.8) or are found in pairs (binary asteroids).

Fig. 4.8 Asteroid *Ida*, and its satellite *Dactyl*. Note the irregular shape of Ida, which is typical for most asteroids and causes variations in the light curve

About 5,000 new objects are discovered per month and more than 300,000 have been registered. For about half of them, the orbit is well established and they have been given permanent numbers. The mass of all asteroids of the Main Belt is estimated to be about $3.0 - 3.6 \times 10^{21}$ kg. This is only about 4% the mass of the Moon (the mass of the Moon is 1/81 that of the Earth). Ceres comprises 1/3 of that mass, 4 Vesta about 9%, 2 Pallas about 7%, and 10 Hygiea about 3%. These are the most massive asteroids. Asteroid 3 Juno has only 0.9 %. The diameter of 1 Ceres (Though it is classified as a dwarf planet, it is mentioned in this chapter because its orbit is typical for Main Belt asteroids.) is between 900 and 1,000 km. The two other large asteroids in the Main Belt are 2 Pallas and 4 Vesta with a diameter of about 500 km. Theoretically, Vesta is visible to the naked eye when it is near to the Earth. The total number of asteroids in the Main Belt is estimated to be $> 10^6$.

4.2.4.2 Classification

Asteroids can be classified according to different criteria.
Classification due to their orbits:

- Main Belt Asteroids,
- Earth crossing Asteroids,
- Trojan Asteroids, etc.

Classification due spectral characteristics based on color and albedo:

- C-Type asteroids — carbonaceous, 75% of known asteroids,
- S-type asteroids — silicaceous, 17% of known asteroids,
- M-type asteroids — metallic, 8% of known asteroids.

4.2.5 Small Solar System Bodies and Habitability

The study of small solar system bodies like comets and asteroids is important for several reasons:

- These objects contain matter as it was present in the primitive solar nebula; the matter was not changed by differentiation processes.
- These objects can be hazardous to life on Earth. There are many Earth crossing asteroids and the orbits of comets are very difficult to determine since their orbits change due to their outgassing processes. We will discuss the impacts of asteroids and comets on Earth and the consequences for habitability in the next chapters.
- Impacts of comets and asteroids on early Earth could have deposited organic compounds and water on our planet, and, therefore, played an important role for the emergence of life.

4.3 Extrasolar Planetary Systems

4.3.1 Some Considerations

The Earth reflects only about 30% of the light received from the Sun (*albedo*, depends on the cloudiness, ice coverage, etc.). The Sun emits 2×10^9 more light than the Earth. However, in the IR at a wavelength of say 10 μm, the Sun-Earth contrast reduces to 10^7. Let us assume that an observer at a distance of 3,000 Lyr wants to observe the solar system. To this observer, the distance Sun-Earth would be only 0.001 arcsec. We can estimate how big a telescope should be in order to resolve this small angular distance: the resolving power of a telescope depends on its diameter D and the wavelength of observation:

$$\alpha = 206265 \times 1.22 \frac{\lambda}{D} \tag{4.11}$$

Now, let $\alpha = 0.001''$ and λ be 500 nm. Then, we see that the telescope aperture should be 125.8 m. To resolve the diameter of the Sun, the aperture should be 150 times as large and, therefore, in the range of 20 km. Presently, such instruments seem unrealistic, but from space, using interferometry, such instruments, could be feasible in the near future. With such instruments, we would be able to resolve such details and to observe directly Earth-like planets in our galactic neighborhood. The value of 3,000 Lyr $\sim$ 1 kpc was taken because, as it will be shown later, there is a HZ in our Galaxy about that extension and in the interval of about 3,000 Lyr we can expect to find most candidates for life.

As it has been shown by the above examples, the direct observation of extrasolar planets is extremely difficult, especially when Earth-like extrasolar planets should be detected. However, there are some indirect methods that already led to reliable results and future space missions that are already under realization will give better knowledge about these possible worlds within the next 10 years. Thus, the field of extrasolar planet search is rapidly expanding.

4.3.2 Detection Methods

The following methods are applied to detect extrasolar planets:

- Astrometry: consider two bodies with masses m_1 and m_2. They will attract each other according to Newton's law of gravity, moving around their center of mass (barycenter):

$$F = G\frac{m_1 m_2}{r^2} \tag{4.12}$$

  The more massive one of the two components, the more the center of mass will be located near to that mass (see Fig. 4.9). Let us consider the center of mass of the system Sun and Jupiter (the most massive planet in the solar system). Since the mass of the Sun is about 1,000 times the mass of Jupiter, the center of mass of such a two body system must be located 1,000 times nearer to the Sun than to Jupiter — in fact it is located just outside the Sun's sphere. Therefore, precise position measurements are required to detect the motion of a star around the center of gravity in case of the presence of other planets.
- Radial velocity or Doppler method: as we have seen above, a star will move under the gravitational attraction of a planet. This motion about the center of mass of the system can be detected by small Doppler shifts of the star's spectral lines. The Doppler shifts vary periodically. This method has been, so far, the most promising one and majority of the extrasolar planets detected up to now were found by these Doppler shifts.
  The principle of radial velocity measurement is illustrated in Fig. 4.10.
  The inclination of the orbital plane with the sky is unknown and often denoted by the letter i. Therefore, we measure the velocity

$$v_r = v \sin i \tag{4.13}$$

  where v denotes the true velocity. When i is equal to 90° we see the orbit edge on. It is evident that small angles of i lead also to low values for the measured radial velocities implying small Doppler shifts. This reduces, dramatically, the detection probability of planets with small i. A further complication is that one has to wait quite a long time to find periodic tiny Doppler shifts. In the case of the

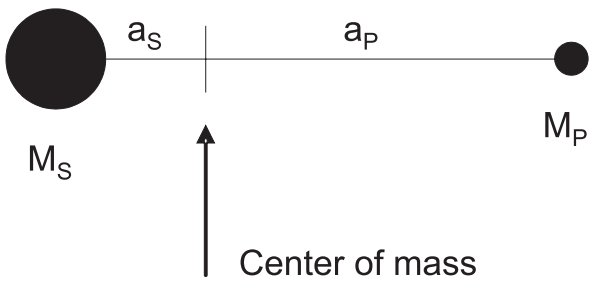

Fig. 4.9 Center of mass in a system of a star with mass M_S and a planet with mass M_P

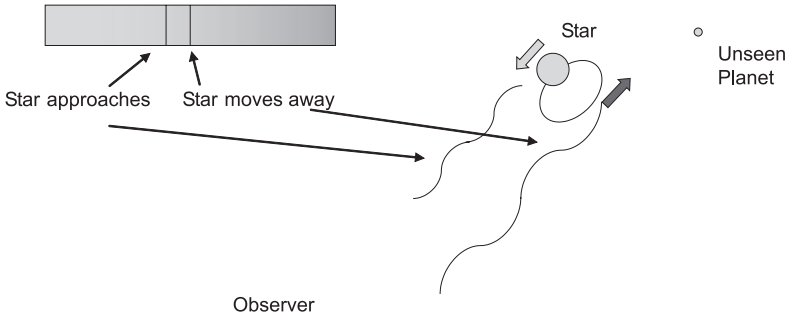

Fig. 4.10 Principle of radial velocity measurement. Note that the spectral lines of the star are blueshifted when it approaches toward the observer and redshifted when its distance to the observer increase

Sun-Jupiter system, the period is 12 years and the Sun's radial velocity around the center of the Sun-Jupiter system is 14 m/s.
When we want to detect planets of Earth masses, then the velocity variations would become about 1/300 of that of Jupiter because the Earth's mass is only 1/300 M_J. It is, therefore, necessary to measure velocities in the range of cm/s. This seems unrealistic but keep in mind that for the Sun, such velocity measurements are already done at regular basis.

- Pulsar timing: pulsars are neutron stars that rotate rapidly. Due to the enormous concentration of the magnetic fields and the inclination of the magnetic field axis to the rotational axis (oblique rotator), particles that are accelerated by the magnetic field radiate like the beam of a light tower. Most pulsars emit radio waves and the pulses can be timed with high precision. Anomalies in the timing can have different causes. The system loses rotational energy due to the acceleration of particles, due to the rapid rotation gravity waves emitted which yields additional loss of energy. Therefore, the older a pulsar, the slower is its rotation rate. These are long-time effects. But anomalies in timing the pulses are also caused by the presence of planets when the pulsar itself moves around the center of gravity, which is a short-time effect.
- Transit method: planetary transits occur only when a planet crosses in front of its parent star's disk seen from the Earth. If the planet is big enough, then a small drop in the brightness of a star can be observed. The decrease in brightness depends on the size of the planet and the size of the star as well as on the distance of the planet from the star. Therefore, it is more probable to observe transits of giant planets that are relatively close to their central star.
  Let us consider the transit method in more detail: Kepler's third law states that

$$P^2 M_* = a^3 \tag{4.14}$$

Given the orbital P of a planet, we can calculate the semi-major axis a of its orbit, when the mass of the star M_* is known. The transit duration time can be found from the formula:

$$\tau_{\rm tr} = 13 d_* \sqrt{a/M_*} \sim 13\sqrt{a}\ \text{hrs} \tag{4.15}$$

Here, d_* is the stellar diameter in solar diameters, M_* the stellar mass in solar masses, and a is in AU. The duration alone does not give any information about the physical nature of the planet. The size of a planet follows from the transit depth because the fractional change in brightness (transit depth) is equal to the ratio of the planet's area to the star's area. As it was pointed out in the chapter about the common properties of stars, the spectral type of a star tells us its size (for main sequence stars).

As an example, we give the transit properties of solar system objects in Table 4.3 from http://kepler.nasa.gov/sci/basis/character.html. P is the orbital period in years, a the semi-major axis in AU, T the transit duration in hours, D the transit depth in %, Prop the geometric probability in %, and Incl the inclination invariant plane in degrees.

Since transits can only be detected if the planetary orbit is near the line-of-sight between observer and the star, it can be derived that the geometric probability for seeing a transit for any random planetary orbit is

$$\sim d_*/2a \tag{4.16}$$

- Gravitational microlensing: microlensing occurs when the gravitational field of a star acts like a lens, magnifying the light of a distant background star. If the foreground lensing star has a planet, then that planet's own gravitational field can make a detectable contribution to the lensing effect. This can be observed in the variation of the lightcurve as a secondary bump (see Fig. 4.12).

  In Fig. 4.11, D_L denotes the distance between the observer O and the lens, D_{LS} is the distance between the lensing star and the source S, and D_S is the distance between the observer and the source.

  An amplification of the light due to microlensing can be observed when the object moves within the *Einstein radius*, which is defined by

$$\Theta_E = \sqrt{\frac{4GM}{c^2}\frac{D_{LS}}{D_L D_S}} \tag{4.17}$$

Table 4.3 Transit properties of solar system planets

Planet	P $P^2 M_* = a^3$	a	T (h) $13\sqrt{a}$	D (%) $(d_p/d_*)^2$	Prop $d*/D$	Incl
Mercury	0.241	0.39	8.1	0.0012	1.19	6.33
Venus	0.615	0.72	11.0	0.0076	0.65	2.16
Earth	1.0	1.0	13.0	0.0084	0.47	1.65
Mars	1.88	1.52	16.0	0.0024	0.31	1.71
Jupiter	11.86	5.2	29.6	1.01	0.089	0.39
Saturn	29.5	9.5	40.1	0.75	0.049	0.87
Uranus	84.0	19.2	57.0	0.135	0.024	1.09
Neptune	164.8	30.1	71.3	0.127	0.015	0.72

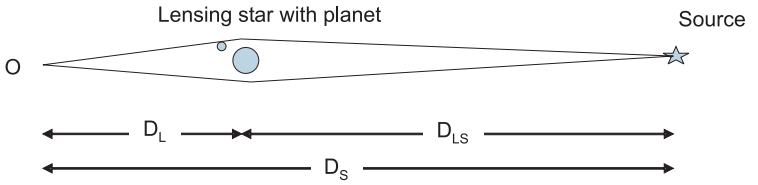

Fig. 4.11 Principle of microlensing. The light of the source *S* is slightly deflected by the lensing star and, additionally, by the planet. *O* denotes the observer

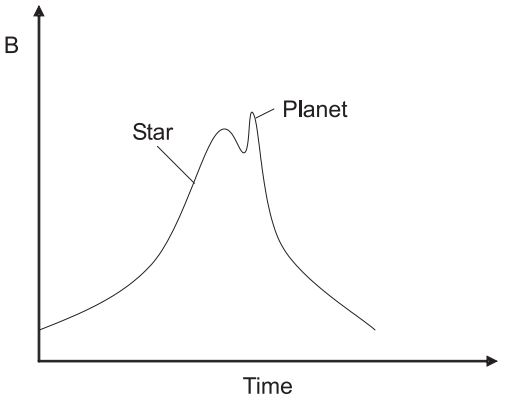

Fig. 4.12 Variation of brightness *B* due to the microlensing effect of a star and its planet

- Light reflected by a planet: it is clear that a planet orbiting a star scatters back into space some amount of the starlight it receives. The ratio of the scattered flux f_P from the planet to the direct flux f_* from the star depends on the size of the planet and its distance to the star. Let us consider a planet of radius R_P in orbit with a semi-major axis a and geometric albedo $p(\lambda)$, then

$$\varepsilon = f_P / f_* = p(\lambda)(R_P/a)^2 \tag{4.18}$$

  The term albedo generally denotes the overall reflection coefficient of an object and determines its brightness. There are two definitions of albedo. The *geometric albedo* is the ratio of actual brightness at zero phase angle to that of an idealized flat, fully reflecting diffusively scattering disk with the same cross section. The *Bond albedo* is the total radiation reflected from an object compared to the total incident radiation from the Sun. Some examples of the geometric albedo are given in Table 4.4.[9]
  The high albedo for Venus comes from the reflection of light at the dense cloud layers of its atmosphere.

[9] From [187].

Table 4.4 Geometric Albedo and Bond Albedo for different objects

Object	Geometric Albedo	Bond Albedo
Mercury	0.138	0.119
Venus	0.84	0.75
Earth	0.367	0.29
Moon	0.113	0.123
Mars	0.15	0.16
Pluto	0.44-0.61	0.4
Jupiter		0.343 ± 0.032
Saturn		0.342 ± 0.030
Uranus		0.290 ± 0.051
Neptune		0.31 ± 0.04

This concept was applied successfully to the star τ *Boo* and its planetary companion. The semi-major axis of the planet is

$$\mathrm{a} = 0.0462(M_*/1.2M_\odot)^{1/3},$$

and its orbital period 3.3 d. If it is a giant planet, with Jupiter-like size and albedo, then $\varepsilon \sim 10^{-4}$. Therefore, the scattered starlight will be at least 10,000 times fainter than the star (under best conditions!). The angular separation between the star and the planet is 0.003 arcsec. Note that the pattern of the photospheric absorption lines in the starlight is preserved when the starlight is reflected from the planet, however, some modifications occur:

– Doppler shift due to the planet's orbital velocity,
– multiplicative scaling by the wavelength albedo of its atmosphere.

Let us assume that the inclination of the orbit is i, and if $K_P = V_P \sin i$, then the Doppler shift varies between $-K_P$ and K_P. The star's absorption lines span $\pm 28\,\mathrm{km/s}$ because of the star's rotation. So, the orbital Doppler shift should be clearly distinguishable, i.e., the planet's spectral line from those of the star. The scattered light should be visible in each spectrum as a faint Doppler shifted copy of each of the star's spectral lines. For more detail, see the paper of Cameron et al. [32].

- Circumstellar disks: there are many cases of circumstellar disks. Their detection and identification is essential for the theories of planet formation. Moreover, there are also cases in which small distortions in the disks have been observed. These distortions can be attributed to planets. The observation of the disks is made preferentially in the IR because the dust absorbs starlight and re-emits it as infrared radiation. As an example, we give, here, the star β *Pictoris* an A5 V star surrounded by a circumstellar disk, which is nearly seen edge on (see Fig. 4.13). This disk was first observed in 1984 and the extension is about 1,000 AU.[10]
  Circumstellar disks were detected near young stars (e.g., the T Tauri stars) or near old stars. It was found that 70–80% of the young stars in the dense clusters of the

[10] The distance from the Sun to the dwarf planet Pluto is only 1/20 of that value.

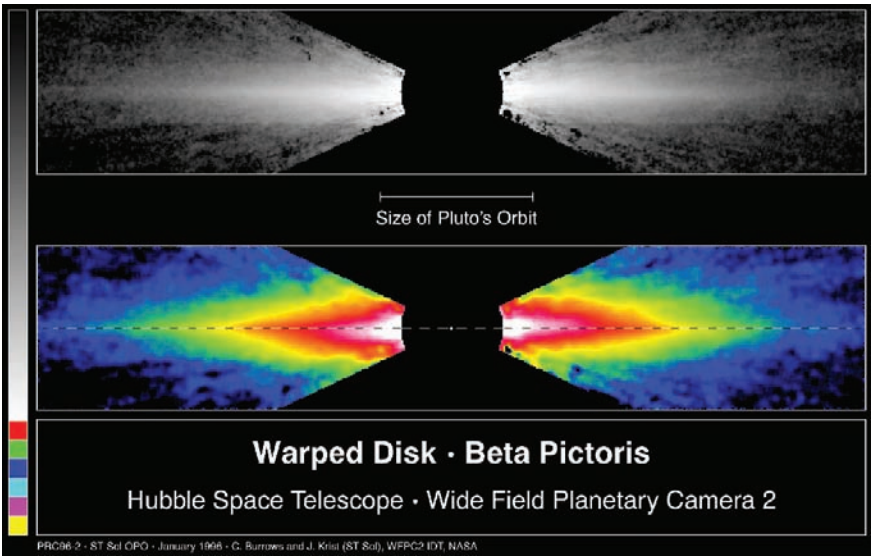

Fig. 4.13 The warped dust disk around β Pictoris. The star itself has been occulted artificially

Orion nebula are surrounded by circumstellar disks, most of them are accretion disks.

The Sun also is surrounded by a disk. This can be seen by the observation of the *zodiacal light* that is concentrated in the orbital plane of the Earth (which is practically the plane of the solar system). The young Sun underwent a T Tauri phase with intense solar wind that blew out all material except larger bodies. Thus, the dust we observe now cannot be primordial dust but it is believed to have been newly formed by collisions of asteroids and evaporation of comets. Observations with the IRAS satellite have shown that 15% of A, F, G, and K main sequence stars have higher infrared excess emission than the Sun (by at least a factor of 100), which indicates a circumstellar dust disk. For our study on the habitability, it has to be stressed that the pure presence of a dust disk does not imply that planets have been formed. But it proves that planet-type objects (asteroids, comets maybe dwarf planets) had been formed during the original star formation process. Otherwise, the percentage cited above must be much less.

Another observation was made by Lagage et al. [133]. They have imaged mid-infrared emission from PAHs at the surface of the disk surrounding the young intermediate-mass star *HD 97048*. It is assumed that this disk is a precursor of more evolved disks like, e.g., that surrounding the star β *Pictoris*.

4.3.3 Examples of Extrasolar Planets

As of May 2008, the following numbers are given:

1. Candidates detected by radial velocity:
   - 235 planetary systems

- 271 planets
- 25 multiplanet systems.

2. Candidates detected by transit:

- 46 planetary systems
- 46 planets
- 0 multiple planet systems.

3. Candidates detected by microlensing: 6 planets.
4. Candidates detected by imaging: 5 planets.
5. Candidates pulsar planets:

- 3 planetary systems
- 5 planets
- 1 multiple planetary system

These data were extracted from the online data catalog of extrasolar planets (maintained by Jean Schneider, CNRS-LUTH, Paris-Meudon Observatory). There are also unconfirmed observations and so-called cluster or free floating candidates for planets.

4.3.3.1 Free Floating Planets

The young nearby stellar cluster around the massive star σ Orionis was studied by optical and IR observations. The cluster is at quite low distance to us (about 350 pc) and extremely young (between 1 and 5 million years). In the paper of Zapatero et al. [272] three objects (S Ori 52, S Ori 60, and S Ori 56) were found with low surface temperatures (derived from optical and near IR low resolution spectroscopy, the temperatures were between 1,700 and 2,200 K). The masses were estimated in the range 5–15 M_J by comparing the observed magnitudes with those predicted by evolutionary tracks. Thus, these objects are unable to maintain nuclear fusion processes to generate energy (their masses are too low). The most remarkable fact, however, was that the objects are not gravitationally bound to any stars. Therefore, they can be regarded as isolated giant planets. From the age of the cluster, it follows that they must have formed on time scales of less than a few 10^6 years.

It could be possible that isolated giant planets form commonly in the universe and that they could significantly populate the galactic disc and even the solar neighborhood. The question is how they could have been formed. The ejection of planetary objects from their orbits around stars is discussed. However, to be consistent with the observations, this process must happen within a relatively short period of a few million years after star formation.

The object *S Ori 70* (mass 3 M_J, temperature 1,000 K, radius 1.6 R_J, and distance 440 pc) was designated by Zapatero-Ososrio et al. [273] as a methane, isolated planetary like one. Another explanation is that it is part of a planetary system surrounding the quintuple star HR 1931 or a foreground field brown dwarf [30]. The object *Cha 110913* (mass 8 M_J, temperature 3000 K, radius 1.8 R_J, and distance 50 pc) is

surrounded by a circumstellar disk [150]. The authors found an IR excess emission at $\lambda > 5\,\mu$m, which can be modelled by an irradiated accretion disk and the accretion rate is supposed $\dot{M} \leq 10^{-12}\,\mathrm{M}_\odot/\mathrm{yr}$. Thus, raw materials for planet formation exist around free-floating bodies with planetary masses.

4.3.3.2 51 Pegasi

The object 51 Pegasi is a Sun-like star of spectral type G2.5V or G4-5Va. Its distance is 50.1 Lyr. It was the first Sun-like star found to have a planet, discovered in 1995 by M. Mayor and D. Queloz at the Observatoire de Haute Provence with the radial velocity method[11] [161], [73] (Fig. 4.14). The parameters of 51 Peg b are as follows:

- Mass: $0.468 \pm 0.007\ \mathrm{M}_J$
- Orbital period: 4.23077 ± 0.00005 days
- Semi-major axis: 0.052 AU (Mercury in the solar system: 0.38!)
- Eccentricity: 0.0

The object is a big Jupiter-like planet that is extremely close to a Sun-like star. The estimated surface temperature is about 1,200°C. Its close position to the central star can only be explained by planet migration.

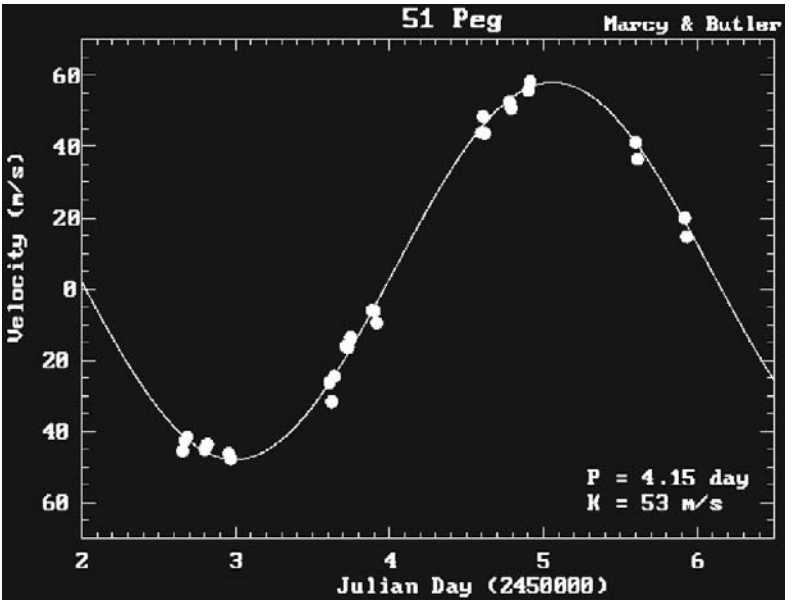

Fig. 4.14 The radial velocity curve of 51 Peg

[11] They used the ELOIDE spectrograph.

The parameters for the central star are as follows: mass $1.06\,M_{\odot}$, Radius $1.15-1.4\,R_{\odot}$, Luminosity $1.30\,L_{\odot}$, effective temperature 5,665 K, Rotation 37 days, and age $7.5-8.5\times10^{9}$ yr.

This object has been studied with the MOST satellite.[12] Since the planet is at a small distance, gravitational interactions with its host star are probable and it has been found that the star shows evidence of activity increases (measured by brightness) induced on the hemisphere facing its planet [256].

4.3.3.3 55 Cancri

55 Cnc is a binary star at a distance of 40 Lyr in the constellation of cancer. It consists of a yellow dwarf star (G8V) and a red dwarf (M3.5-4V). The separation between the two stars is 1,000 AU. The system further contains four planetary objects, three of which are similar to Jupiter and one is similar to Neptune. The parameters for the four planets are given in Table 4.5.

This object has been studied by Gonzales [77], where other objects can be found (51 Peg, 47 UMa, 70 Vir, and HD 114762). A fifth planet was detected, with an apparent orbital period of 260 days, placing it 0.78 AU from the star in the large empty zone between two other planets [60].

4.3.3.4 Gliese 581c

The star Gliese 581 was already known as a star with planets and its distance is 20.5 Lyr. A Neptune-mass planet was detected and there is strong evidence for a planet with eight Earth masses. The object Gliese 581c (in the constellation of *Libra*) was detected by European astronomers[13] and is probably a planet with a diameter of 1.5 times the Earth's diameter [252]. It completes a full orbit in 13 days. This object is one of the smallest and most earth-like object found so far. Its distance to the central star is only about 10 million km (which is much less than the distance of mercury to the Sun). However, one must take into account that Gliese 581 is a red dwarf star. Its luminosity is 50 times fainter than the Sun. Therefore, the conditions on Gliese 581c

Table 4.5 Parameters of the planetary system 55 Cancri

Object	Mass (M_J)	Orbital Period (d)	Sem. Axis (AU)	Ecc.
e	$> 0.045 \pm 0.01$	2.81 ± 0.002	0.038 ± 0.001	0.174 ± 0.127
b	$> 0.784 \pm 0.09$	14.67 ± 0.0006	0.115 ± 0.003	0.0197 ± 0.01
c	$> 0.217 \pm 0.04$	43.03 ± 0.021	0.240 ± 0.005	0.44 ± 0.08
d	$> 3.92 \pm 0.5$	4517.4 ± 77.8	5.258 ± 0.9	0.327 ± 0.28

[12] Microvariability & Oscillations of Stars.

[13] See ESO Science Release 222/07.

could be quite favorable and it may be located in a HZ. It is estimated that the mean surface temperatures could lie in the interval between 0 and 40°C. This implies that water would be liquid. From the radius of 1.5 Earth radii, it can also be estimated that this "Super Earth" has a rocky surface similar to our Earth and possibly large parts covered by oceans. Thus, Gliese 581c is an important target for future research because it is quite nearby.

In Fig. 4.15, a three-planet Keplerian model of the Gliese 581 radial-velocity variations is given. The panels display the phase-folded curve of each of the planets, with points representing the observed radial velocities, after removing the effect of the other planets. Top panel refers to the 15 Earth-mass planet orbiting close to the star (5 d period), the middle one is the 5 Earth-mass planet in the HZ and the lower

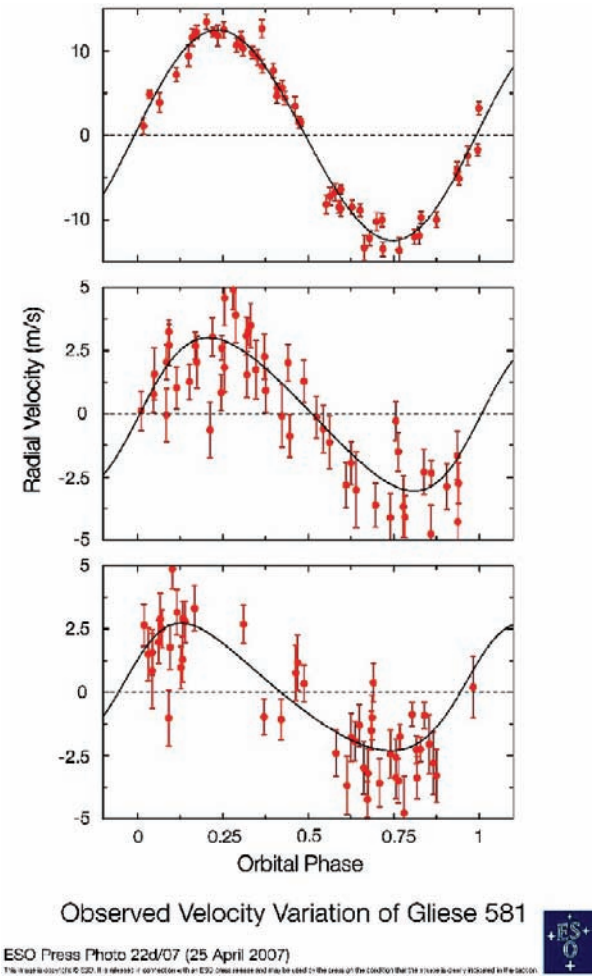

Fig. 4.15 The Gliese 581 planetary system

panel shows evidence for a third 8 Earth-mass planet with a period of 84 days. The error on one measurement is of the order of 1 m/s.

The detection of Gliese 581c was made by using a very sophisticated instrument at a 3.6 m telescope at La Silla, Chile: The High Accuracy Radial Velocity Planetary Searcher (HARPS).[14] HARPS (Fig. 4.16) measures the motions of the central star that are expected due to the gravitational pull of the planets. With that instrument, radial velocities on the order of 1 m/s can be detected. HARPS is an echelle spectrograph and contained in a vacuum vessel to avoid spectral drift due to temperature or air pressure variations. There are two fibre cells which feed the spectrograph; one collects star light, the other is used to record simultaneously a Th-Ar reference spectrum or the background sky. The aperture of the two HARPS fibres is $1''$ on the sky— a resolving power of the spectrograph of 115, 000 is obtained. The observations can be made in the spectral range of 378 nm − 691 nm in the echelle orders between 89 and 161. The detector consists of two CCDs (in total 4k $\times$ 4k). To give an example, let us consider the following numbers: at 550 nm, the signal to noise ratio is 110 for a sixth magnitude G2V star in 1 minute integration time ($1''$ seeing). With the simultaneous Th-Reference, the signal to noise ratio should be sufficient to determine radial velocities of about 0.90 m/s.

Fig. 4.16 The HARPS tank. Courtesy: ESA

[14] See also: http://www.ls.eso.org/lasilla/sciops/3p6/harps/

4.3.4 Planet Formation around Pulsars

The first extrasolar planetary system was discovered by *A. Wolszczan* around the pulsar *PSR B1257+12* in 1992, 2 years after the Pulsar itself was found with the Arecibo Radio Telescope [270]. The objects were detected by variations in the pulse arrival times due to the reflex motion of the neutron star induced by the planets. Any object near such a pulsar would cause a motion about the center of gravity of the pulsar itself and this motion causes tiny shifts in the arrival times of the pulses. The variations were about $\pm 1.5\,\mathrm{ms}$. This corresponds to a $\pm 0.7\,\mathrm{m/s}$ variation in the pulsar's radial velocity. Even asteroid masses could be detected by such measurements.

Two Earth sized planets orbit the pulsar *B1257+12* and a third body of the mass of our Earthmoon has been detected in this system. The main parameters are given in Table 4.6. The pulsar is in the Virgo constellation at a distance of 900 light years and its age is estimated to about 800 million years. We might speculate that the intense radiation beam of the pulsar might not cover the planets because of their highly inclined orbits. Lazio et al. [139] discuss the search for exoplanets around pulsars.

Shabanova [224] found a planet with at least two Earth masses and a semi-major axis of 7.3 AU orbiting the Pulsar *PSR B0329+54*. This is a very slow pulsar, that means a relatively old object (pulsation period about 0.7 s).

In all of the above cited cases, gas disks with a mass of $\sim 0.1\,\mathrm{M}_\odot$ form. For more details, see the paper of Hansen [90]. With the Spitzer Space telescope a circumstellar disk around the Pulsar *4U 0142+61* was found. This object is in the constellation *Cassiopeia* about 13,000 Lyr away and may have formed 100,000 yr ago. The disk of debris from the supernova explosion orbits at a distance of about 2 million km and contains about ten Earth masses. Its presence can be measured by its IR radiation [257].

4.3.4.1 Formation of Pulsar Planets

The main question is whether these planets survived the violent supernova explosion or have been formed from the remnants of such an explosion.

Table 4.6 Confirmed Pulsar Planets (2007), adapted from Gozdziewski, Konacki, and Wolszcan [79]

Pulsar	Planet	Mass
PSR B1620-26	PSR B1620-26c	2.5 M_J
PSR 1257+12		
	PSR 1257+12 a	0.020 M_E
	PSR 1257+12 b	4.3 M_E
	PSR 1257+12 c	3.9 M_E
	PSR 1257+12 d	0.0004 M_E

There are several explanations. Circumpolar accretion disks could explain both planets and spin down of pulsars. Such an accretion may have been formed by mass transfer in a binary system, in which one companion evaporated. The evaporation of a companion object in the very late phases of binary millisecond pulsar evolution or low mass X-ray binaries was discussed by Banit et al. [8]. Generally, the evaporation of a companion star may take up to 10^8 y, which is in agreement to the timescales found. Another more catastrophic scenario is the collision of a binary companion with the neutron star during the supernova explosion; material from the tidal disruption of the companion provided material for the disk. This scenario reminds to the theory of Chamberlain, 1901, who claimed that the solar system was formed by strong tidal interaction during a close encounter of the Sun with a neighboring star. Such a scenario for the formation of ordinary planetary systems is very unlikely because, then, planetary systems would be very rare and the Deuterium problem cannot be solved. We measure certain amounts of Deuterium on the planets. If they were of solar origin, then, almost no Deuterium should be present, since this was consumed in the very early phases of solar evolution by nuclear fusion.

4.4 Stability of Planetary Systems

Unstable orbits of planets strongly influence the habitability of these objects. In this section, we discuss these problems, but first, a short introduction to the definition of stability is given.

4.4.1 What does Stability Mean?

Two hundred years ago, the solar system was considered a stable system in which no changes occur. *Newton* formulated the law of gravitation and the laws of motion. From that time, the question arose: do we live in a deterministic universe in which all future motions, accelerations, actions, and interactions are determined by the present position and motion of all the particles? Such a point of view would imply to negate the notion of free will for humans. Examples of deterministic systems are typical topics in classical physics such as Newton's law of motion, classical electrodynamics, and thermodynamics. The point is, if the initial conditions were known in enough detail, the future evolution of the system is fully predictable. This leads us to *Laplace's demon*.[15] Imagine an entity (today, we would say a supercomputer) that knows all facts about the past and present, that means it knows exactly the initial states defined by the positions and velocities of all particles in the universe. If all forces are known, then, it would be possible to predict entirely the future.

If predictability cannot be made for all future, then, this could be because there is a lack of information. Applied to our problem, this would mean that we should try to measure as accurate as possible the positions and velocities of the objects in the

[15] 1814, Pierre-Simon Laplace.

solar system. There is also a *biological determinism*: all behavior, belief, and desire are fixed by our genetic environment. Again, the problem of free will is unsolved.

Examples of nondeterministic systems are completely random events.

In *quantum physics*, exactly which atom will undergo a radioactive decay cannot be predicted. The *Schrödinger cat* is a very famous thought experiment. Suppose a cat is in a steel chamber. In a Geiger counter, there is a radioactive substance. This substance should have the probability that in the course of 1 hour, one of the atoms decays but with equal probability that it will not decay. If it decays, the counter tube discharges and through a relay, a hammer is released which destroys a flask containing hydrocyanic acid, which would kill the cat. The quantum mechanical Ψ function gives for the probability of the whole system 0.5. Thus, after 1 hour, the cat is in a quantum superposition of coexisting living and dead states. As soon as the box is opened, only one of the states, "dead" or "alive", is observed, not a mixture of them. Thus, in quantum physics, predictability is reduced to probability.

Let us go back to the problem of stability of planetary orbits. In celestial mechanics, this reduces to the solution of the *n-body problem*. This problem cannot be solved analytically for $n > 2$. Therefore, we must use some numerical procedures to solve the equations which are always approximations. Furthermore, we certainly do not know with ultimate precision all the initial state vectors of all the bodies in the solar system. This introduces additional uncertainties. The question is how these uncertainties influence the evolution of the orbits. Could small random perturbation lead to completely different orbits that may even become unstable. In this context, we have to define "chaos".

For a dynamical system to be classified as *chaotic*, the following properties can be listed:

- Sensitivity to initial conditions: this is also known as the "butterfly effect" (E. Lorenz, 1972: does the Flap of a butterfly's wings in Brazil set off a Tornado in Texas?). An extremely small change in the initial condition of a system causes a completely different result.
- Topologically mixing: the system will evolve over time. Any given region of its phase space will eventually overlap with any other given region. The phase space of a dynamical system consists of all coordinates and momentum variables.
- The periodic orbits must be dense. There exist dynamical systems that are chaotic everywhere (Anosov diffeomorphism) – however in many cases chaotic behavior is found in a subset of phase space – it occurs on an attractor. That means a large set of initial conditions will lead to orbits that converge to this attractor. There are simple attractors (e.g. points or circle like curves, limit cycles). Chaotic motion gives rise to strange attractors.

An example is the *Lorenz attractor*,[16] which describes convective rolls in the atmosphere and the system becomes chaotic for a certain set of parameters.

A measure of how a system evolves from two infinitesimally close states is the *Lyapunov exponent*. Let us consider two trajectories in phase space with initial separation $\delta \mathbf{Z}_0$. Then, two trajectories in phase space diverge according to

[16] Lorenz, 1963.

$$|\delta \mathbf{Z}(t)| \sim e^{\lambda t} |\delta \mathbf{Z}_0| \tag{4.19}$$

One usually can search for the largest Lyapunov exponent. If this is positive, then, two points that are initially separated by an infinitesimal small amount will diverge exponentially. The interpretation is simple: if there is a system with positive Lyapunov exponents, then, an infinitesimal small error in the initial conditions will yield to completely different evolution.

The maximum Lyapunov exponent follows from

$$\lambda = \lim_{t \to \infty} \frac{1}{t} \log \frac{|\delta \mathbf{Z}(t)|}{|\delta \mathbf{Z}_0|} \tag{4.20}$$

If a system is *conservative*, a volume element in phase space stays along a trajectory and the sum of all Lyapunov exponents must be zero. If a system is *dissipative*, the sum of Lyapunov exponents is negative.

These methods of modern chaos theory can be applied to investigate whether catastrophic changes of planetary orbits are likely or not.

4.4.2 The Earth: Perturbations by the Moon and Planets

The Earth is in an elliptical orbit around the Sun and it defines a two-dimensional plane which is called the *ecliptic*. Some basic data are given in Table 4.7.

Table 4.7 Characteristic data for the Earth's orbit

Aphelion distance	152.097701×10^9 m
Perihelion distance	147.098072×10^9 m
Semi-major axis, a	149.5978875×10^9 m
Semi.minor axis, b	140.5769998×10^9 m
Eccentricity e	0.016710219
Sidereal period	365.256366 d
Orbital speed: average	29.783 km/s
Orbital speed: maximum	30.287 km/s
Orbital speed: minimum	29.291 km/s

The orbital elements of a planet are given in Fig. 4.17.

There are six orbital elements:

- Semi-major axis a
- Eccentricity of the orbit e. From this value, we can calculate the distance when a planet is closest to the Sun (perihelion)

$$r_{\text{Perihelion}} = a(1-e) \tag{4.21}$$

and the distance when the planet is at its greatest distance to the Sun:

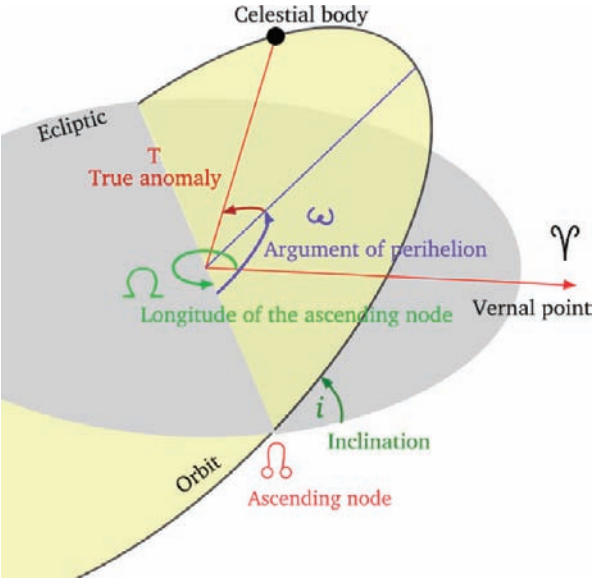

Fig. 4.17 Orbital elements. The vernal point is defined as the location of the position of the Sun at the beginning of spring. Courtesy: Arpad Horvath

$$r_{\text{Aphelion}} = a(1+e) \tag{4.22}$$

- i Inclination of the orbit,
- ω longitude of the perihelion,
- Ω longitude of the ascending node (planet passes from below to above the ecliptic).

Because the orbital inclinations are small in the solar system, also, the angle

$$\bar{\Omega} = \omega + \Omega$$

is used.

The solar system can be described as a 10 body system: The Sun + 8 planets + Pluto (dwarf planet). Due to the mutual perturbations, the orbits of the planets will change. For the orbital parameters of the Earth, the following variations are known:

- *Precession* of the Earth's axis: the Earth rotates around its axis (spin ω_S about spin axis), this leads to an angular velocity of rotation. Because the axis of rotation is not perpendicular to its orbit, and, because, due to its rotation, the Earth is slightly flattened, the Moon (being at most $\pm 6^\circ$ north or south of the ecliptic) and the Sun (being per definition in the ecliptic) exert a torque Q on it. This results into the angular velocity of precession:

$$\omega_P = \frac{Q}{I_S \omega_S} \tag{4.23}$$

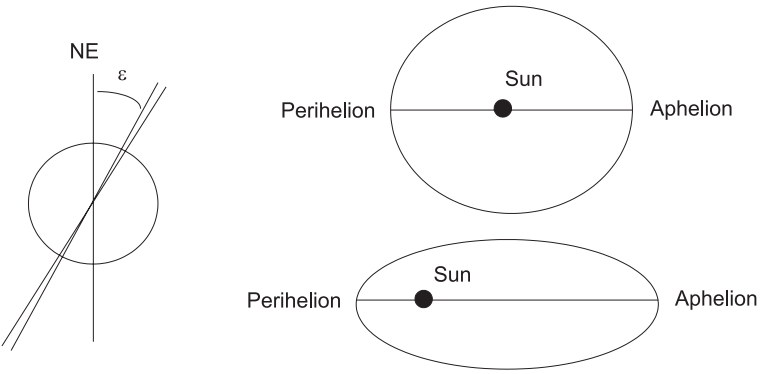

Fig. 4.18 *Left*: change of the Earth's rotational axis tilt; ε is the angle between the axis of rotation and the normal to the ecliptic (NE ecliptic north pole); *right*: effect of the eccentricity e on the perihelion and aphelion distance

I_S is the moment of inertia. Using $\omega = 2\pi/T$, the period of precession is

$$T_P = \frac{4\pi^2 I_S}{QT_S} \tag{4.24}$$

The Earth's poles complete a circle around the ecliptic pole (which is normal to the plane of the ecliptic) over a period of 25,800 years. At present, there is the star *Polaris* about 1° away from the north-celestial pole. The precession is 1° in 180 years and was detected by the Greek astronomer *Hipparchos*. In fact, we have to distinct the precession caused by the Sun and the Moon, lunisolar precession, from the precession caused by the planets. Due to this effect, the perihelion changes within the orbital plane.

- Change of the *ellipticity* of the Earth's orbit: the period of this change is between 90,000 and 100,000 years. When the orbit is highly elliptical, the differences in the amount of solar insolation between perihelion and aphelion increase to more than 20% (Fig. 4.18).
- Change of the *obliquity* of the Earth's axis: the cycle for the change of the tilt, which varies between 22.1 and 24.5 degrees, is about 40,000 years. More tilt means that seasons are more pronounced: colder winters and hotter summers. Less tilt means that the differences between the seasons become smaller, cooler summers and milder winters (Fig. 4.18).

4.4.2.1 Ice Ages

According to *Milankovich*,[17] these variations are the main cause for the *ice ages*, which show a periodicity of about 100,000 years. If we take the orbital variations

[17] Milutin Milankovitch, 1879–1958.

[194] and the solar forcing (this was calculated for 1st of July solar insulation at a latitude of 65°N using Jonathan Lewine's insolation calculator and combine them with glacial data [146], then, we arrive at Fig. 4.19. The interglacial periods are shown as grey and defined as deviations in the 5 kyr averaged of at least 0.8 standard deviations above the mean. The solar radiation received during summer at 65°N latitude is critical for the growth and decay of ice sheets and was taken in many cases as a prediction of the world climate curve.

The theory of Milankovitch was first completely ignored. In 1976, deep-sea sediment cores were examined for past climate variations. Hays *et al.*, 1976 [97] extracted the temperature change up to 450,000 BP.[18] They found that these variations are in good agreement with the above mentioned changes of the Earth's eccentricity, obliquity, and precession. Up to this study, it was believed that, e.g., external factors like variations in the output of the Sun, variations in the amount of energy of the Sun reaching the Earth caused by changing interstellar dust concentrations, volcanic dust, changes in the Earth's magnetic field or factors like distribution of carbon dioxide in the atmosphere and deep ocean circulation might be responsible for the explanation of the Pleistocene ice ages. Astronomical causes for the onset of ice ages are discussed by Muller [174].

The climate reconstruction was made by three measurements:

- δ^{18}O isotopic composition of planktonic foraminifera. This isotope has two more neutrons than the common ^{16}O. If it is bound in water, then, this water molecule is heavier. Thus, cooler water releases more ^{16}O than ^{18}O. Cooler air precipitates more ^{18}O than warmer air. Cooler water collects more ^{18}O than warmer water and, thus, from the abundance of ^{18}O the temperature can be estimated.
- Analysis of radiolarian assemblages.
- Percentage of Cycladophora davisiana.

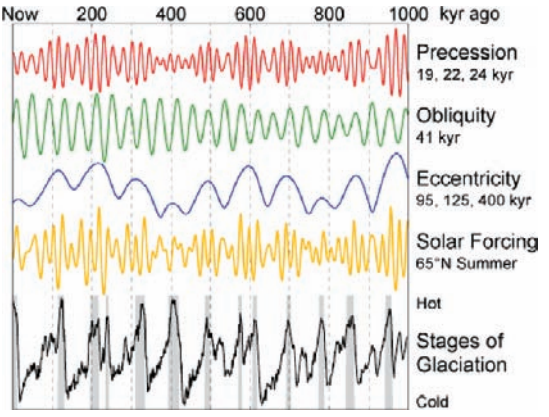

Fig. 4.19 The Milankovitch cycles. Warm periods are shaded gray. Note that the increase to a warm period is less steep than the evolution toward a cold period. Courtesy: Global Warming Art

[18] BP is used for *before present*.

The end of an ice age could also be triggered by large volcanic eruptions. Also, the distribution of the continents (which changes because of plate tectonics) plays an important role because the flow of warm water from the equator to the poles is affected. Ice sheets increase the albedo of the Earth, which means a positive feedback. The flow of warm water to the poles is reduced when, e.g., a continent is there (Antarctic today) or when there is a large continent about the equator. The most recent North American glaciation was the Wisconsin glaciation (70,000–10,000 yr BP). The ice sheets extended to about 45° north latitudes and were 3 to 4 km thick. The lowering of the sea level was about 120 m. Ice ages in Europe are listed in Table 4.8.

Table 4.8 Ice Ages in Europe

Name	Period
Würm	120,000–10,000 BP
Riß	240,000–180,000 BP
Mindel	480,000–430,000 BP
Günz	640,000–540,000 BP

4.4.2.2 Stability of the Earth's Orbit and Rotation Axis

The stability of the Earth's orbit was tested by a 3-million-year integration [194].

Another question which is of great importance for climate variations on Earth is the *stability of the Earth's obliquity*. Laskar et al. [138] performed an integration of the equations of precession of the Earth. They found a chaotic zone between 60°and 90° in obliquity. However, because of the torque exerted by the *Moon*, the obliquity of the Earth is stable and the variations are only $\pm 1.3^\circ$ around a mean value of 23.3°. If the Moon was not present, two effects would happen:

- The chaotic zone would extend from 0° to 85°.
- This would induce catastrophic climate changes on the planet, maybe destroy its habitability.
- The Earth would have a faster spin rate. This would lead to a greater flattening and maybe some compensation because of the larger action of the torque of the Sun and the planets on the more flattened Earth.

Quasiperiodic variations in the obliquity of Earth in the absence of the Moon have been also suggested by Ward [258]. Because of the tidal evolution of the Earth-Moon system, the days will become longer and the lunar orbit expands. Ward also shows that in less than 2×10^9 yr, the equinoctial precession periods will become 49,000 years when the semi-major axis of the moon's orbit will be $66.5\,R_{earth}$ and 69,000 years when the semi-major axis of the moon's orbit becomes $68.0\,R_{earth}$. At the same time, this will give rise to large oscillations of the Earth's obliquity and, therefore, to large climate fluctuations destroying the Earth's habitability.

4.4.3 Is the Solar System Stable?

To test the stability of the solar system, that means to answer the question of whether strong perturbations or even collisions between planets are probable or whether planets can be expelled out of the solar system, long term integrations of the n-body problem have been performed by different groups. The equations of motion in a heliospheric coordinate system for a given planet with mass m_i are given as follows:

$$\frac{d^2\mathbf{r}_i}{dt^2} = -G\frac{M_\odot + m_i}{r_i^3}\mathbf{r}_i + \sum_{j=1, j\neq i}^{9} \frac{Gm_j}{|\mathbf{r}_i - \mathbf{r}_j|^3}(\mathbf{r}_j - \mathbf{r}_i) - \sum_{j=1, j\neq i}^{9} \frac{Gm_j}{\mathbf{r}_j^3}\mathbf{r}_j \qquad (4.25)$$

Here, $\mathbf{r}_i, \mathbf{r}_j$ denote the coordinates of the planets i, j; m_i, m_j their masses, and $M_\odot$ the mass of the Sun. Note that the first term on the right side describes a two body problem (two bodies with masses $M_\odot$ and m_i attract each other) and the other terms are perturbations by the other planets.

The accuracy of the integration depends on

- Accuracy of initial positions $\mathbf{r}_i(t_0)$,
- Accuracy of initial velocities $\dot{\mathbf{r}}_i(t_0)$,
- Accuracy of initial masses m_i.

It is also relevant for such a problem to treat the Earth–Moon system in an appropriate way. One way is to replace Earth and Moon by circular rings of masses m_{earth} and m_{moon} and radii R_{earth} and R_{moon}. These are located in the plane of the ecliptic and centered on the Earth–Moon barycenter. The main conclusion of that work was that all planetary orbits appear stable and do not show exponential divergence during that period [194].

Birn [20] studied the stability of the planetary system by numerical integration starting with various initial conditions and found that stable orbits occur only within certain stability zones. The present orbits of the planets are found at the central part of these stability zones.

Very long numerical integrations of the solar system (8 planets + dwarf planet Pluto) were performed by Ito and Tanikawa [115]. The time span of the integration was 10^9 years. The behavior of the planets appeared quite stable with no apparent secular increases of eccentricities or inclinations of the planets. However, a potentially diffuse character of the motion of Mercury was found. This result is similar to that found by Laskar. The maximum eccentricity was found there to be 0.35 over ± 4 Gyr. Laskar detected in 1989 [136] that using a numerical integration over 200 million years, the inner planets of the solar system showed chaotic behavior of their orbital motion. In 1994, he performed a numerical integration of the eight planets 10 Gyr backward and 15 Gyr forward. He found that the large planets Jupiter, Saturn, Uranus, and Neptune behave quite regularly. The curves of the variations of the eccentricity and inclination for the inner planets, however, exhibit a diffusive behavior

(see Fig. 1 in [137]). *Mercury's* eccentricity can potentially reach values very close to 1. This would be caused by a close encounter with Venus in less than 3.5 Gyr. Thus, Mercury is likely to be ejected out of the solar system from a close encounter with our sister planet Venus.

4.4.4 Pluto—Charon and Triton

As we have mentioned, *Pluto* is no longer considered to belong to the family of planets but to be a member of so-called dwarf planets. The diameter of Pluto is $2,274$ km and its mass is 1.27×10^{22} kg. The dwarf planet was discovered in 1930 by *C. W. Tombaugh* at Lowell Observatory.

In 1978, Pluto's largest satellite, *Charon* was detected. From the distance and the orbital period around Pluto, the mass of Charon (1.9×10^{21} kg) can be easily derived: 1/7 th the mass of Pluto. Charon's diameter is 1,172 km and the average separation from Pluto is just 19,640 km. Due to this small distance, Charon's orbit is gravitationally locked with Pluto. Both bodies keep the same hemisphere facing each other. The rotational periods of Pluto and Charon as well as Charon's orbital period are 6.387 days.

The density of Charon is between 1.2 and $1.3\,\mathrm{g/cm^3}$ and Pluto's density is 1.8 to $2.1\,\mathrm{g/cm^3}$. Thus, both objects should have formed independently and they cannot have formed together by accretion as a double dwarf planet.

For our considerations about catastrophes and stability, it is interesting to mention that Charon could have been formed when a planetesimal impacted on Pluto, similar to the theory about the formation of the Earth's moon. A hint for such a scenario is the relatively high inclination of Pluto's rotation axis and Charon's orbit (98.8°).

Another theory states that Charon, Pluto, and Neptune's satellite Triton (Fig. 4.20) belong to the icy outer solar system and were not swept up by Uranus and Neptune or ejected from the solar system. These three bodies have several similarities.

In 2005, Hubble Space Telescope observations showed two other tiny moons of Pluto, namely Nix and Hydra with diameters between 60 and 200 km. Pluto is locked in a 3:2 resonance with Neptune that means that his orbital period is exactly 1.5 times longer than Neptune's. Because of the high inclination of its orbit, Pluto will not collide with Neptune, even though because of the large eccentricity of Pluto's orbit, the dwarf planet is sometimes nearer to the Sun than Neptune.

Triton is the largest satellite of Neptune: diameter 2,700 km, mass 2.14×10^{22} kg. It is the only large moon in the solar system, that orbits retrograde.[19] It seems now well established that Triton must have been formed at a different region and not in the vicinity of Neptune, maybe in the *Kuiper belt* and later it was captured by Neptune. Because of its retrograde orbit, Triton is slowed down by

[19] Some moons of Jupiter like Ananke, Carme, Pasiphae, and Sinope and of Saturn like Phoebe also move retrograde but their diameters are less than 1/10 the diameter of Triton.

Fig. 4.20 Triton, 1989, Voyager 2 flyby. The pinkish deposits constitute a vast south polar cap believed to contain methane ice, which would have reacted under sunlight to form pink or red compounds. The dark streaks overlying these pink ices are believed an icy and perhaps carbonaceous dust deposited from huge geyser-like plumes, some of which were found active during the Voyager 2 flyby. The bluish-green band visible in this image extends all the way around Triton near the equator; it may consist of relatively fresh nitrogen frost deposits. The greenish areas includes what is called the cataloupe terrain, whose origin is unknown, and a set of "cryovolcanic" landscapes apparently produced by icy-cold liquids (now frozen) erupted from Triton's interior. Courtesy: Calvin J. Hamilton

Neptune's strong tidal interaction. Therefore, its orbit becomes closer to Neptune and it could either break up or form a ring. The rotation axis of Triton is tilted 157° with respect to Neptune's. Neptune's rotation axis is inclined 30° from the plane of its orbit. Thus, similar to Uranus, the poles of Triton are exposed to the Sun. The density of Triton is $2.0\,\mathrm{g/cm^3}$ and it is composed about 25% of water ice. The satellite has a very thin atmosphere (0.01 mbar) composed mostly of nitrogen and a small amount of methane. The surface temperature is similar to that of Pluto: 34.5 K. Since only few craters are visible, the surface must be very young, consisting of frozen nitrogen and methane. It was also found that Triton has ice volcanoes.

4.5 Extrasolar Planetary Systems

Having discussed the process of the formation of our planetary system, let us now investigate the extrasolar planetary systems so far detected and how these findings are in accordance with what we suppose to know from the solar system formation.

4.5.1 Stability of Orbits in Binary Stars

4.5.1.1 Restricted Three Body Problem

In this section, we will give a short overview of a special problem in celestial mechanics that has many applications in extrasolar planet research as well as in solar system calculations: the *circular restricted three body problem*. We assume two massive bodies (e.g., a double star) and investigate the motion of a third particle in that system. The mass of the third particle should be negligible, not influencing the motions of the primary bodies. The two massive bodies are assumed in circular orbits around their common center of mass.

In such a system, there are five points, the so called *Lagrangian points*, where the third body can be stationary relative to the two primaries (Fig. 4.21). Calculations show that the points L1, L2, and L3 are unstable. They lie at the line connecting the two primaries. Small perturbations lead to instability. The points L4 and L5 are at the apex of two equilateral triangles formed by the points m_1, m_2, L_4 and m_1, m_2, L_5. These are stable points. Perturbations (not too large) lead to a libration of the m_3 mass around L_4 resp. L_5. Stability holds as long as the mass ratio between the two large masses exceeds 24.96.

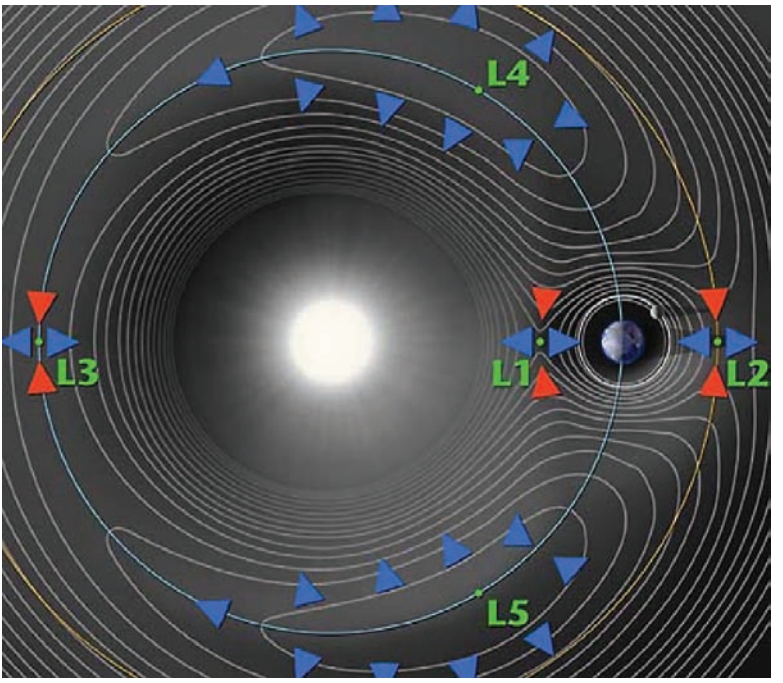

Fig. 4.21 Lagrangian points in the Sun-Earth system. At L1, the solar satellite SOHO is located (pointing always to the Sun); at L2, the James Webb telescope will be placed in 2011

In the solar system, the Trojan Asteroids are found at L_4, L_5 in the Sun-Jupiter system.

4.5.1.2 Extrasolar Planets in Double Stars

There are many double and multiple stars (about 50% of all stars). For extrasolar planet research, it would be interesting to know whether stable orbits of planets can exist in multiple systems. For a system of two stars, the answer is similar to the restricted three body problem. A possible planet can be either near m_1 or near m_2 or farther away orbiting both primaries. Raghavan et al. [197] discuss this problem.

Let us give two examples that illustrates that stable planetary orbits in binary stars are possible. The system α *Cen* A and B is the nearest system to the Sun. The luminosity of α Cen A is $1.5\,L_\odot$, α Cen B is fainter with $0.45\,L_\odot$. Their motion about each other is in an eccentric orbit with the extreme distances 35 AU and 11 AU. However, there exists a region where HZ about both stars can exist. This zone is about 3 AU about each star. Another well-known example of a binary with a planet is *Gliese 86* . This is a spectroscopic binary star. A planet was found in this system at 0.08 AU distance from the star. Since the primary is a cool dwarf, the surface temperature on the planet is 380°C and its orbital period 15.83 days. The separation of the two binaries is assumed 100 times the distance between the planet and the primary, thus a stable orbit is possible (see Queloz et al. [193].

4.5.2 Migration of Planets

The process of formation of the planetary system described above must be refined because of new observational facts. From the classical theory, it is hard to understand why there are Jupiter-sized planets found close near their host stars. This can be explained by migration of planets during the formation phase of a planetary system.

The giant planets in the solar system are gas planets and they must have been formed in a gas-rich environment. The gas was acrreted from a protoplanetary disk. Strong tidal interactions between a protoplanet and the disk may induce gap formation in the disk and, thus, the protoplanet's mass growth is stopped because there is no mass in the gap. The results of Lin and Papaloizou [144] show

- If the protoplanet is initially located at small distances from the protosun, it would migrate inward.
- If the protoplanet is initially at large distances, it would migrate outward until the disk matter interior to its orbit becomes sufficiently depleted.

Thus, the migration of planets during the early stages of planetary formation was found. Applying these results to our solar system, the authors concluded that Jupiter must have formed just outside the present orbit of Mars. The mass of the disk

exterior to the orbit of Jupiter must have been less than 0.1 $M_\odot$, otherwise, Jupiter would have spiralled toward the protosun in less than 10^6 yr.

Planet migration can be understood by a slight imbalance of the gravity field created by the spiral density waves [259]. When the gravitational force exerted by the outer disk is greater than the one created by the inner disk, a planet looses angular momentum and spirals inward. Without further assumptions, this spiralling would be very fast (some possible braking mechanisms are the effect of the magnetic field, torques at corotation, interactions between several planets, etc.). There are two types of migration:

- Type I: no gap forms, the net torque exerted on a planet by the disk is zero. If the disk is thick, then, the protoplanet drifts on a time scale inversely proportional to its mass.
- Type II: the protoplanet opens a gap, the planet migrates on a time scale that is set by the disk's viscosity.

Type I migration seems to occur on much faster time scales than type II.

These theoretical considerations were confirmed a decade later by the discovery of Jupiter-sized planets at distances less than 0.4 AU from their host stars.[20]

The migration and growth of protoplanets in protostellar disks was studied by Nelson et al. [177]. They found that the direction of orbital migration is always inward and that the protoplanet finally has a near circular orbit after a characteristic viscous time scale of 10^4 initial orbital periods about its central star. There seems to exist also a limit of about 5 M_J for closely orbiting giant planets.

The growth and dynamical evolution of protoplanetary material from planetesimals to earth-like planets during and after the migration of a giant planet through the terrestrial zone was studied by Raymond et al. [204]. They simulated a disk of 17 Earth masses extending from 0.25 to 10 AU and containing 80 Moon to-Mars sized protoplanets and 1,200 planetesimals. Furthermore, they assumed a gradient in the disk's composition, the inner disk being Fe-rich and water-poor, 50% water beyond 5 AU. Then, a Jupiter-mass giant planet was assumed to start at 5 AU and migrate within 0.25 AU. The integration, time of the whole system was 200×10^6 yr. At the end of the integration, the system has evolved to inner disk composed of planetesimals at 0.06 AU; at 0.12 AU; a planet with 4 M_E; a hot Jupiter at 0.21 AU; and a planet with 3 M_E at 0.91 AU. Such a system can be stable over $\sim 10^9$ years. This is the time needed for evolution of life on Earth.

Dynamical interactions in multiplanet systems were studied by Marzari and Weidenschilling [157]. Three Jupiter-mass planets were assumed to orbit a one solar-mass star. After 10^8 years, some of the tested configurations became chaotic. Due to close encounters, one object was ejected of the system, one moved closer to the star, and the third remained at its position. The innermost orbit of the planet corresponded to about 1/2 of its starting semi-major axis. Thus, by gravitational scattering the existence of hot Jupiters, Jupiter-sized planets very close to their host stars, cannot be explained.

[20] 0.4 AU is about the distance of Mercury from the Sun.

Chapter 5
Catastrophes in Our Solar System?

In this chapter we first discuss possible particle and radiation hazards that could destroy habitability on a planet. Then, collisions with comets and asteroids will be investigated. We mention, here, only catastrophes that can occur within the solar system; influences from outside the solar system will be discussed in the next chapter.

5.1 Catastrophes by Particles and Radiation Hazards

5.1.1 Major Solar Events

5.1.1.1 Flares

As we have discussed in the chapter about the Sun, solar flares release huge amounts of energy within a short timescale. Flares produce a burst of radiation in nearly the whole electromagnetic spectrum but mainly in X-ray, gamma ray, and EUV as well as in the radio range (Fig. 5.1). Therefore, flares are classified according to their X-ray brightness in the wavelength range from 1 to 8 Å[1]:

- X-class flares, $I > 10^{-4}\,\mathrm{W/m^2}$: major events, trigger planet-wide radio blackouts and long lasting radiation storms. The short wavelength radiation influences on the ionization of the upper atmosphere of the Earth.
- M-class flares, $10^{-5} < I < 10^{-4}$: medium sized. Can cause brief radio blackouts affecting mainly the Earth's polar regions.
- C-class flares $10^{-6} < I < 10^{-5}$: small flares, no consequences on Earth.

One of the most powerful events was the so called *Bastille Day* Event that occurred on 14 July 2000 (Figs. 5.2 and 5.3). About 15 min after the observed eruption, energetic protons from the flare arrived at Earth and they triggered a S3 radiation storm event. Such an event can cause single-event upsets on satellites, noise in imaging systems, permanent damage to exposed components/detectors, and decrease in

[1] 1Å = 0.1 nm.

A. Hanslmeier, *Habitability and Cosmic Catastrophes*, Advances in Astrobiology and Biogeophysics, DOI 10.1007/978-3-540-76945-3_5,

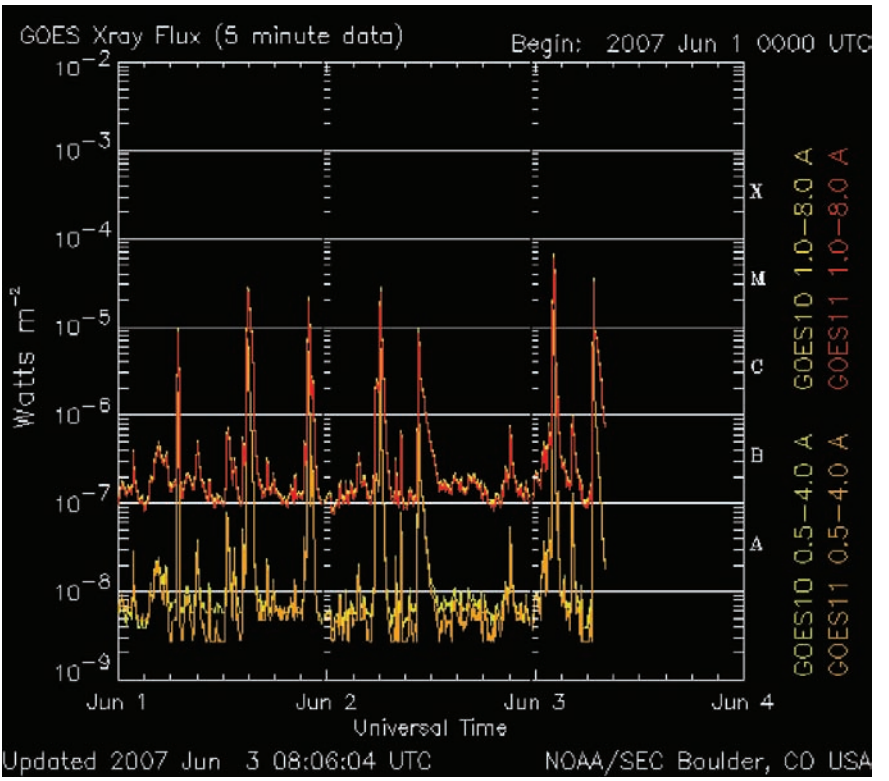

Fig. 5.1 The solar X-ray flux is constantly monitored by the GOES satellite. Large X-ray bursts cause short wave fades for HF propagation paths through the sunlit hemisphere. Some large flares are accompanied by strong solar radio bursts that may interfere with satellite downlinks

solar panel currents. On Earth, air travellers at high latitude are exposed to radiation that is equivalent of a brief chest X-ray. The flare occurred at 10.24 UT and the SOHO satellite observed a full-halo coronal mass ejection. It was estimated that 10^9 t of solar plasma was ejected and the speed toward Earth was between 1,300 and 1,800 km/s. These particles hit the magnetosphere where they are deflected. The magnetosphere becomes compressed and *magnetic storms* evolved. Such storms induce electric currents that can interfere with electric power transmission on the surface. Also aurorae are produced, which can then be seen at low magnetic latitudes.

The global response of the low-latitude to mid-latitude ionosphere due to the Bastille Day flare was studied by Huba et al. [109].

Other famous solar major events were the *Halloween events* in October/November 2003.

5.1.1.2 Ionosphere

The ionosphere is the uppermost part of the Earth's atmosphere and it is ionized by the shortwave solar radiation and, therefore, strongly influenced by any variation of

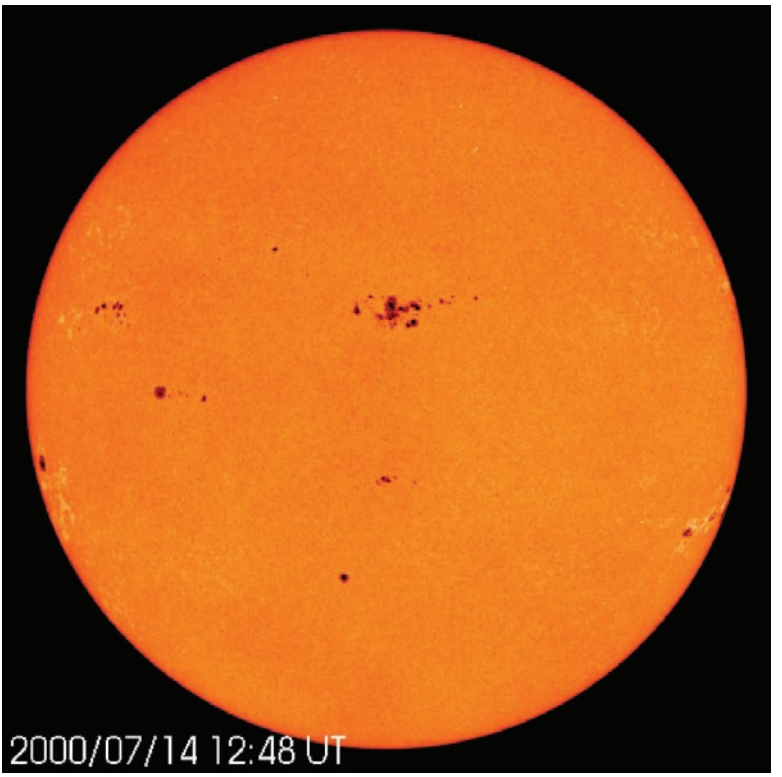

Fig. 5.2 The Sun with the sunspot group that produced the solar flare, Bastille Day Event. Photo: SOHO/MDI, ESA/NASA

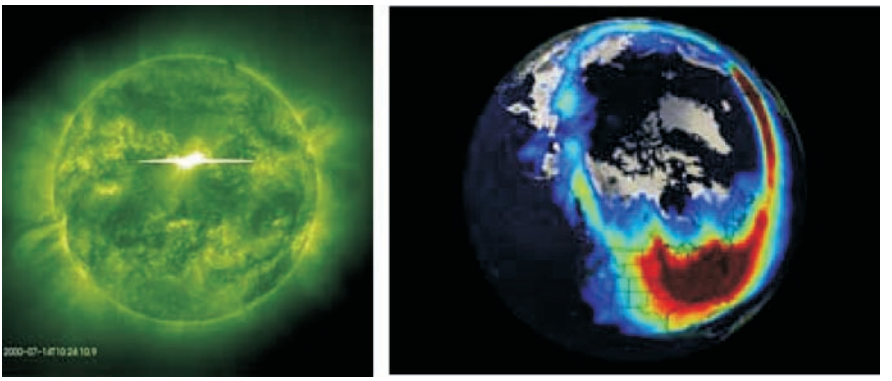

Fig. 5.3 Bastille day flare. On the ground, aurorae were seen as far south as El Paso, Texas. Power companies suffered geomagnetically induced currents that tripped capacitors and damaged at least one transformer. Global positioning system (GPS) accuracy was degraded for several hours. The image on the right side shows in false colors the geomagnetic storm that was induced 31 hours after the solar event on Earth. Photo: SOHO/MDI, Polar, ESA/NASA

the amount of that input. In the deeper layers of the Earth's atmosphere, there is a balance between ionization and recombination because the density is large enough that free electrons get captured.

The ionosphere consists of several layers (Fig. 5.4):

- The D-layer is between 50 and 90 km, ionization of NO due to the hydrogen Lyman α (121.5 nm), ionization of N_2, O_2 by hard X-rays. During night, cosmic rays produce a residual ionization. The recombination is high (because of density) and, therefore, the ionization is low. HF radio waves are not reflected in this layer (absorption of HF below 10 MHz). During daytime, the distant AM broadcast band is absorbed.
- E-layer: between 90 and 120 km, ionization of molecular O_2 due to solar soft X-ray (1–10 nm) and far UV radiation. Reflects radio frequencies < 10 MHz. During night this layer disappears because the ionization source is not present.
- E_s is the sporadic E-layer, where there are small clouds of intense ionization, and radio waves are reflected in the band 25–225 MHz. These events may last from minutes to hours.
- F-layer: 120–400 km. Solar EUV radiation (10–100 nm) ionizes atomic oxygen (O). For HF communication, this is the most important part of the ionosphere; during daytime it splits into F_1 and F_2. The F-layers are responsible for skywave propagation of radio waves. These layers are the most reflective layers of radio waves on the side of the Earth facing the Sun.

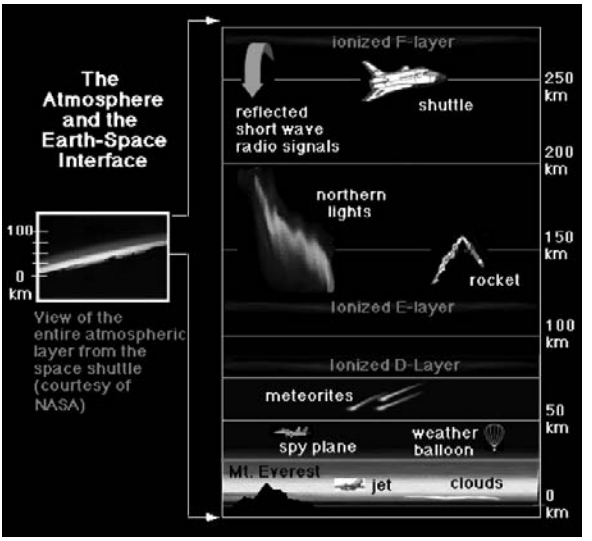

Fig. 5.4 Atmospheric layers. Courtesy of Windows to the Universe, http://www.windows.ucar.edu

5.1.1.3 Solar Flares and Disturbances in the Earth's Atmosphere

In the Earth's ionosphere, the *total electron content (TEC)* is an important parameter; the number of free electrons is determined by the ionization. The amount of ionization depends on solar radiation, which changes because of a diurnal effect, a seasonal effect (in winter one hemisphere receives less radiation), and solar activity in general. *Sudden ionospheric disturbances (SID)* occur on the sunlit side of the Earth and are caused by hard X-rays that are emitted from strong solar flares. They penetrate to the D-region, increase ionization there, and, thus, the number of electrons increases. This causes increased absorption in the 3–30 MHz range and a radio blackout. On the other hand, however, very low frequency signals (3–30 KHz) become reflected by the D-layer.[2]

Solar flares also release high-energy protons. These can reach the Earth within 15 min to 2 hours after the onset of a solar flare. The particles (protons) spiral down the magnetic field lines of the Earth and penetrate into the atmosphere near the magnetic poles increasing the ionization of the D and E layers there. The so-called *Polar Cap Absorption (PCA)* lasts from hours to days (on the average between 24 and 36 hours).

During a *geomagnetic storm*, the F_2 layers becomes unstable and may disappear completely. In the polar regions, *aurora* will be observable.

The *critical frequency* is the limiting frequency below which a radio wave is reflected back by an ionospheric layer[3]:

$$f_{\text{crit}} = 9 \times 10^{-3} \sqrt{N_e} \tag{5.1}$$

where N_e is the electron density per cm^3 and the frequency f_{crit} is given in MHz.

The state of the magnetosphere is given by several indices. The *A* and *K* indices measure the behavior of the horizontal component of the geomagnetic field (K index is from 0 to 9).

The Halloween event on 28 October 2003 caused a 30 % increase in the local noon equatorial ionospheric column density. This is usually measured by a TECU (total electron content unit) and

$$1\,\text{TECU} = 10^{16} \text{electrons}/\text{m}^2 \text{ column density} \tag{5.2}$$

The Halloween event caused a 25 TECU increase within 5 min, which was five times the increases for the Bastille Day Event (14 July 2000) and the events for 29 October, Nov. 2003. Also, in the 26–34 nm EUV range, the 28 October peak was more than twice as large as for the other three events [247]. A ring current is built up as magnetic reconnection between the southwardly directed interplanetary magnetic field and the Earth's magnetic field occurs.

---

[2] Normally these frequencies are reflected by the E-layer but are damped in the D-layer.

[3] At vertical incidence, otherwise, the formula given must be divided by $\sin\alpha$, where α is the angle of attack.

The ozone depletion during a solar proton precipitation in the October 2003 events cited above was calculated by a model [56].

The enhancement of NO_x during the large solar proton events in Oct./Nov. 2003 and the resulting ozone depletion was studied by Lopez-Puertas et al. [148], and Rohen et al. [213]. They claim that the ozone loss is dominated by

- Above 50 km: HO_x such as H, OH, and H_2O.
- Below 50 km: NO_x such as NO and NO_2.

From the GOES-11 satellite data, the ion electron pair, odd-nitrogen, and odd-hydrogen production was estimated and it was found that the odd-hydrogen causes the ozone depletion by a factor of 2 at altitudes higher than 55 km, which is also in agreement with measurements.

5.1.1.4 Effects on Habitability

A decrease of stratospheric ozone will result in an increase of UV radiation at the surface. Mutagenic effects of solar UV-radiation on DNA were studied, e.g., by Rabbow and Horneck [195]. They used a sample of the plasmid DNA pUC19 and its *E. coli* host JM83 and exposed this sample to radiation between 280 and 400 nm wavelength (this corresponds the UV-range under ozone depletion) and 300–400 nm (this simulates the UV-range at present). Their conclusion was that a depletion of stratospheric ozone leads to a considerable increase of mutations and skin cancer, but also the longer wavelength UVA part should be taken into account (315–400 nm). These wavelengths can penetrate proliferating tissues of the skin; mutations may then lead to tumor induction. It is estimated that a 1% loss of stratospheric ozone results in a 2% increase in UV radiation on the Earth's surface. This could result in about a million extra human skin cancers per year [48].

Ionizing radiation-induced mutation of human cells with different DNA repair capacities was investigated by Amundson and Chen [3]

Thus, solar flares, influence on the state of the ionosphere trigger geomagnetic disturbances and even change the composition of the atmosphere above the troposphere by ozone depletion. The activity of the present day Sun, however, does not seem to occur at an amplitude that could cause catastrophic events that would severely affect habitability on Earth. During times of magnetic field reversal also, the effect of enhanced radiation on the surface of the Earth should be taken into account.

Enhanced short wavelength radiation on the Earth's surface increases mutation and can, therefore, also be seen as an important driver for the appearance of new species.

5.1.2 The T Tauri Phase of the Early Sun

Generally, T Tauri stars have deep convection zones and rotate very rapidly. Since the stellar dynamo is caused by the interaction of rotation and convection, their energetic flaring activity can be explained. T Tauri stars are surrounded by an accretion disk. The T Tauri stage lasts only for about $3–4 \times 10^6$ yr. During the T Tauri phase, jets occur at the polar regions. These jets eject about 2% of the infalling matter from the accretion disk and brake down the high rotation rates of the protostars. In the later stages of T Tauri evolution, the accretion slows and the jets decline, however, a strong stellar wind becomes important. This wind sweeps out the remaining gas and dust from the solar system and only planetesimals as well as the Kuiper belt objects remain there.

The early Sun was different from the present Sun in many ways. First, the luminosity was only about 70% of its present-day value. But that does not mean that the early Sun was an inactive star. On the contrary, during its T Tauri phase, very energetic flares occurred [70]. The X-ray intensity of these flares was 100 times more than that of today. Moreover, such high-energetic flares occurred at a frequency of about 1,000 times the present high-energetic flare rate.

Discussing this high rate of super-energetic flares in the context of habitability on early earth, we must also consider that there was no protective ozone layer in the Earth's atmosphere.

Therefore, the T-Tauri Phase of the Sun was important for expelling remaining gas and dust from the inner solar system. On the other hand, no HZ existed in the solar system because of the intense and frequent UV- and X-ray outbursts.

Classical T Tauri stars as future solar systems are discussed by Bertout, Basri, and Cabrit [15].

5.2 Catastrophes in the Early Solar System

5.2.1 Planetesimals

At the end of its T Tauri phase, the intense solar wind has ejected all gas and dust from the disk leaving about 10^{12} planetesimals of the size of about 1 km. For these planetesimals, the radioactive isotope ^{26}Al is of interest. Its half-life is about 720,000 years. Where did this isotope come from? Three possible explanations for this isotope are discussed:

- Close nearby supernova explosion,
- Nearby massive stars at their asymptotic giant branch evolution, strong winds bring material from the stellar cores,
- Produced by high-energy flare events of the protostar.

The importance of this isotope is that during the decay of ^{26}Al, heat was produced. Small planetesimals and dust could very effectively radiate this heat into space, however, for objects larger than several 10^2 km, this heat produced a melting. This caused a *differentiation*, heavier metals like Fe sank down due to gravity and the lighter elements formed a crust. A silicate mantle and a liquid iron core developed for the inner planets of the solar system.

The mass of the dwarf planet *Ceres* is about 1/6,000 that of the Earth. Therefore, 6,000 planetesimals of the size of Ceres (about 1,000 km in diameter) would be needed to form the Earth; certainly, this number was much higher because most bodies were considerably less massive than Ceres. It can be estimated that it took 30–40 million years to accumulate an Earth-sized body.

5.2.2 Heavy Bombardment Phase

The next phase lasted longer and is also known as heavy bombardment phase: from 4.56×10^9 to 3.9×10^9 BP.

The area of heavy bombardment in the solar system is documented by the cratered surfaces of all terrestrial planets and many satellites. The crater sizes follow a power law: every time the diameter is halved, there are eight times more craters. Between 4.1 and 3.8 billion years, a sharp decrease of the heavy bombardment was found (from the manned Apollo missions). This picture was refined in the last years: because of rapid planetary formation from accretion, between 4.4 and 4.0 billion years BP, the impact rate was low and, then, a second peak can be observed around 3.9 billion years BP. This peak can be explained by impactors that came from collisions in the asteroid belt. Thus, there were two major phases of heavy bombardment and catastrophic collisions during the evolution of the solar system:

- During planetary accretion $> 4.4 \times 10^9$ yr BP.
- Late heavy bombardment: maximum around 3.9×10^9 yr BP.

In the first few tens of million years, Ceres-, Moon-, and even Mars-size planetesimals collided. Later, impact of larger bodies may have caused a complete evaporation of the oceans on the early Earth.

At the beginning of the heavy bombardment phase, a crucial event occurred: the proto-Earth collided with a Mars-size planetesimal, which lead to the formation of the Moon.

5.2.3 The Formation of the Moon

There are several facts that any theory of the formation of the Moon has to explain: its density is half that of the Earth; it does not possess an iron core; and it has exactly the same oxygen isotope composition as the Earth.

Hartmann and Davis [96] therefore, explained the formation of the Moon by a collision of the Earth with a Mars-size planetesimal about 4.527×10^9 BP. The two metallic cores (that of the proto-Earth and that of the Mars-size impacting planetesimal) merged and the Moon was formed from the material in the lighter Earth's mantle. The Moon was formed at a distance of about 22,000 km ($3.6\,R_{Earth}$) from the Earth's center.[4] Thus, strong tidal interactions occurred. At the time of the formation of the Moon, the Earth's rotation rate was 5 h and was slowed down by the moon's strong tidal forces.

The size of the impacting body follows from the abundances of Vanadium, Chromium, and Manganese because these elements are depleted in the mantles of the Moon and the Earth and the depletion factors are very similar for the two bodies [118].

The numerical simulations of Ida, Canup, and Steward [112] showed that the Moon could have been formed from accretion in a circumterrestrial disk of debris generated by such a giant impact on Earth in less than 1 year. Canup [33] showed that the collision must have occurred with an impact angle of 45 degrees and an impactor velocity at infinity $< 4\,km/s$

The formation of the Moon may have been crucial for a continuously habitable Earth because due its (the Moon's) presence, the Moon

- Stabilizes the tilt of the Earth's rotation axis.
- Tides occur – it is believed that the tides played an important role in the transition from life in the oceans to life on land.
- The rotation of the Earth slowed down.

5.2.4 Collisions in the Early Solar System

5.2.4.1 Impacts on the Moon

Collisions with large bodies were quite frequent in the early solar system. In the paper by Safronov and Zvjagina [214], it was first pointed out that collisions with large objects were common in the early solar system. This can be deduced, also, from different structures on the lunar surface. The *Mare Imbrium* has a diameter of 1,300 km and is the second largest mare on the moon (the largest is the Oceanus Procellarum). It is surrounded by a ring of mountains that were uplifted by the impact. Some mountains and craters also indicate an inner ring of about 600 km diameter. The initial crater due to the impact could have been 100 km deep. The impact occurred 3.8 to 3.9×10^9 yr ago. After the impact, it took about 100 million years to fill the basin with lava. The kinetic energy of the impact was about 10^{27} J.

The *Mare Orientale* is another large impact bassin, however, it is difficult to see from Earth. The diameter is about 900 km and it is believed to be slightly younger than the Mare Imbrium.

[4] The distance Earth-center – Moon-center is now 63 R_{Earth}.

Let us compare the impact probabilities on Earth and Moon: the mass of the Moon is 1/81 that of Earth, therefore, the impact energy on Earth is larger by that factor. Moreover, the cross section of the Earth is higher (13 times that of the Moon). We must assume that during the early phases, large impacts also occurred on the Earth and that these impacts lead to ocean evaporation.

5.2.4.2 Evaporation of Early Oceans

It can be estimated that a collision with a body of 500 km and mass larger than 1/10 that of *Ceres* could have vaporized a global ocean. Sleep and Zahnle [228] showed that impacting energies $> 10^{28}$ J were large enough to vaporize the oceans. The atmosphere left was hot (2,000°C) and dense (100 atm). This caused the oceans to vaporize within a few months. The atmosphere was extremely heavy but could not escape into space. For about 100 years, the temperature on the Earth's surface remained about 1,500°C. This had catastrophic consequences for possible lifeforms. The Earth changed into an extremely hostile environment. The outmost parts of the atmosphere slowly cooled and due to atmospheric convection also the surface cooled. Within 2,000 years, the oceans attained their old depth again.

Rocks with traces of life or organic compounds could have re-seed the Earth.

The relation between impact energy and water evaporated from a global ocean is roughly given in Table 5.1. From that table it is seen that the *K-T event* that caused the extinction of the Dinosaurs 65 million years before present (BP) released an energy of 10^{24} J and would have been able to lower the global ocean by 30 cm. Some other pronounced impact events on the Moon are also listed such as the formation of the craters *Langrenus*, *Copernicus*, and *Tycho*.

General features of terrestrial planet growth by collisions with planetesimals were investigated by Wetherill [264].

Let us give some other catastrophic examples of collisions in the early solar system.

Table 5.1 Impact energy and water evaporation. Event denotes an impact that occurred on the Moon or Earth; time denotes when that impact happened

Impact Energy (J)	Water evaporated (m)	Event	Time BP
2×10^{35}	3,000	Early Earth	$> 4 \times 10^9$ yr
$\sim 10^{27}$	30	Mare Imbrium (Moon)	$\sim 4.0 \times 10^9$
10^{23}	0.03	Langrenus (Moon)	$\sim 2.0 \times 10^9$
10^{22}	0.003	Copernicus (Moon)	$\sim 900 \times 10^6$
10^{22}	0.003	Tycho (Moon)	$\sim 100 \times 10^6$
10^{24}	0.3	K/T event (Earth)	65×10^6

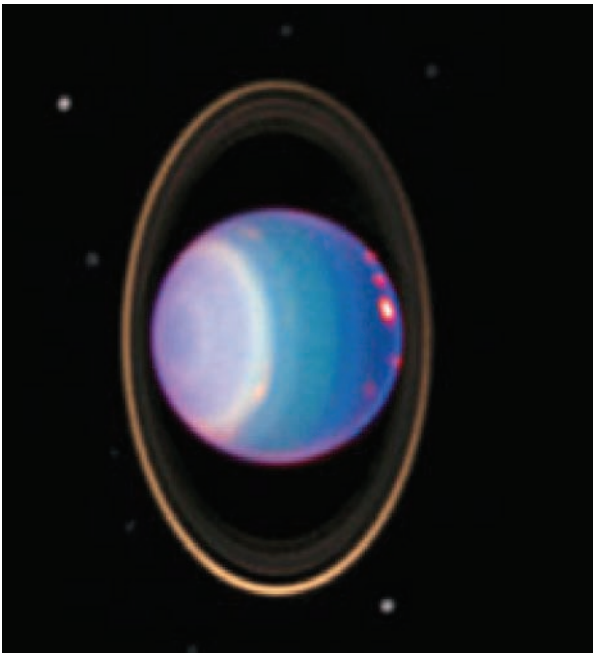

Fig. 5.5 Uranus with its ring system and three satellites. Photo credit: HST

5.2.4.3 Uranus

Current theories about the formation of the planetary system imply that the rotational axis of the planets are perpendicular to their orbital planes. Larger inclinations of the rotational axes have to be explained by collisions.

The obliquity of the axis of Uranus is 98° (Fig. 5.5). The poles point East–West. The north pole aims at the Sun at the beginning of northern summer and away from the Sun at the beginning of southern summer. The orbital period of Uranus is 84.011 years; the northern and southern hemispheres of Uranus are 42 years in sunlight and 42 years in darkness.

To explain this strong deviation, Uranus must have collided with a large object in its early phase. The size of the impacting body could have been 1/10 of the mass of Uranus (which is 14 M_E) (Harris and Ward [92], in that paper, other calculations are also presented).

5.2.4.4 Venus: Retrograde Rotation

One typical property of the Solar System is that all planets (and most smaller objects) orbit the Sun counterclockwise as seen from a position above the Sun's north pole and rotate in this sense except Venus, whose rotation axis has a tilt of less than 3° and a very slow rotation of 243 days and retrograde. This can be explained by

an early massive impact. Studies by Correia and Laskar [44] have shown that this strange behavior may also be a result from a spin axis flip (due to a core-mantle friction inside the planet and atmospheric tides) and that Venus could have started with a prograde rotation and then developed a retrograde rotation for which the obliquity went to zero.

There is a strong tidal interaction between the Sun and Venus and the only explanation why there is no 1:1 rotation to orbital motion resonance is the perturbation by the Earth.

In fact, the evolution of the rotation of Venus seems to depend on several factors:

- Tides from the Sun.
- Tides from the Earth (much weaker),
- Friction with the atmosphere and atmospheric tides. This is transferred to the crust and produces a torque. In the atmosphere of Venus, a *4-day superrotation* is observed; the clouds in the atmosphere rotate much faster than the surface. This superrotation can be explained by meridional circulation and planetary scale waves.

5.3 Collisions in the Solar System

5.3.1 Case Study: Meteor Crater, Arizona

On Earth, a meteor impact occurred 50,000 years ago, a crater 180 m deep with a rim rising 30–60 m above the surrounding plain and a diameter of 1,200 m was formed. This crater is known as the *Meteor crater* or as *Barringer Crater* (see Fig. 5.6). Kring [130] calculated magnitudes of pressures and wind velocities as a function of distance for that event. He assumed two cases for the energy that was set free: 20 megaton and 40 megaton. For comparison, the Hiroshima bomb had only 15 kt equivalent energy, thus, the energy released through the impact corresponded to at least 1,000 Hiroshima bombs. The results are given in Table 5.2.

It can be roughly estimated that the devastated area around Meteor Crater was about 800–1500 km^2. Let us take as a mean value 1,000 km^2 and estimate the chances of it being in these 1,000 km^2 compared to the total surface of the Earth which is 510 $\times 10^6$ km^2. This is 1:500,000. It also can be estimated that the probability of such an impact is 1:1,500, which means that it can occur on an average every 1,500 years. Then, the combined probability for a person standing in the right area at the right time, that is, being killed by an impact that lead to the Meteor Crater is 1 in 7.5×10^8.

In Fig. 5.7, the approximate frequency of impacts vs. megatons TNT equivalent energy is given: for example, in every decade, an event with an approximate equivalent energy of ~ 1 megatons has to be expected. The impact effects of the Meteor Crater event affected only about 1,000 km^2 and no global extinction resulted. From this Figure, it can be deduced that an impact of 10^4 megatons TNT equivalent would

Fig. 5.6 The 50 000 year old Meteor Crater in Arizona

Table 5.2 Magnitudes of pressures and wind velocities as a function of distance for the Meteor Crater impact event [130]

			20 megatons	40 megatons
Peak Overpressure (psi)	Peak Dynamic Pressure (psi)	Max. Wind Velocity (km/h)	Distance (km)	Distance (km)
100	120	2,300	2.8	3.6
50	40	1,500	3.8	4.8
30	17	1,100	4.9	6.2
20	8.1	800	5.9	7.4
10	2.2	470	8.5	11
5	0.6	260	12	16
2	0.1	110	21	27
1	0.02	60	32	40

lead to a global catastrophe. The marked events (Tunguska and K-T impact) will be discussed below.

Let us do another example for the calculation of the probability that a person could be killed due to an impact event.[5] A 10^5 megaton impact certainly would lead to a global catastrophe and such an event could be expected every 5×10^5 years. Let us assume that one in four people would be killed in such a global catastrophe. Then, the chances for any person dying in such an event during the next year is one in two million.[6] The diameter of a crater that such an object would cause is between 10 and 20 km.

[5] After Lind M.V. Martel http://solarsystem.nasa.gov

[6] The chances of being killed in a car accident is about 1 in 5,000.

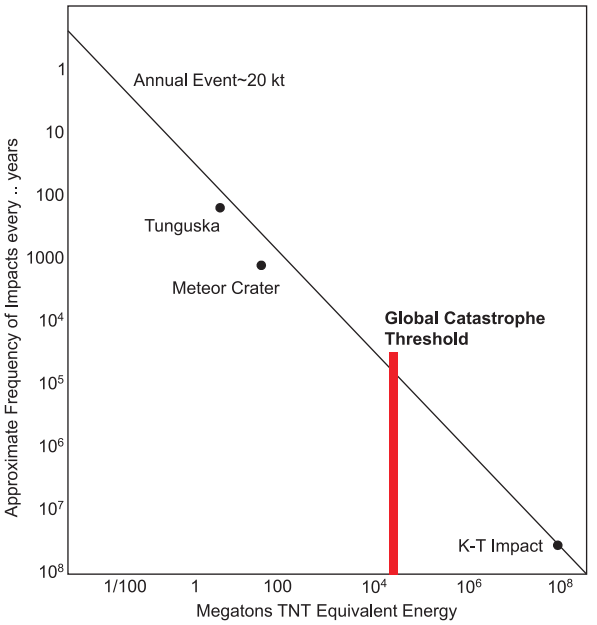

Fig. 5.7 Approximate frequency of impacts vs. megatons TNT equivalent energy (adapted from [130])

5.3.2 The K-T Event

5.3.2.1 Evidence for Impact Theory

Sixty-Five million years ago, about 70% of all species on Earth disappeared within a very short time. This mass extinction is known as the K-T event because it occurred at the Cretaceous-Tertiary border in the Earth's history. In a layer near *Gubbio*, Italy, Alvarez, Asaro, and Michel [1] found a peculiar sedimentary clay layer (only 1 cm thick) that was deposited at the time of a mass extinction event; moreover, this layer contained anomalous amounts of the rare element *iridium*. It can be easily explained why that element is rare on Earth. At the early stage of Earth formation, heavy elements like iridium, platinum, or iron sank down to the core when the Earth was largely molten (*differentiation process*, see planetary formation). Such a differentiation process did not occur on the small bodies of the solar system such as meteoroids or asteroids. These objects still have the primordial solar system composition. To explain the Ir abundance and other element anomalies found in that layer, the impact of a 10 km chondritic asteroid would have been sufficient.[7] Such a large impact would have had approximately the force of 100 trillion tons of TNT, i.e., about 2 million times as great as the most powerful thermonuclear bomb ever tested.

[7] Under the assumption that it contained the normal percentage of iridium found in chondrites.

The impact theory can also be traced back to M. W. DeLaubenfels' hypothesis [50].

Summarizing, the Alvarez impact theory is supported by several observational facts:

- Chondritic meteorites and asteroids contain a much higher iridium concentration than the earth's crust because they have about the same concentration of iridium as the whole earth and were not differentiated.
- The isotopic composition of iridium in asteroids is similar to that of the K-T boundary layer but differs from that of iridium in the earth's crust.
- Chromium isotopic anomalies found in Cretaceous-Tertiary boundary sediments also strongly support the impact theory and suggest that the impact object must have been an asteroid or a comet composed of material similar to carbonaceous chondrites.
- Shocked quartz granules, glass spherules, and tektites are common, especially in deposits found in the Caribbean area.
- All of these constituents are embedded in a layer of clay, which can be interpreted as the debris spread all over the Earth's surface by the impact.

While the element anomaly was the first evidence of an impact event that led to the mass extinction, the question remained as to where on Earth this impact did occur. In 1990, Hildebrand and Boyton [102] became aware of data by geophysicists that were searching for oil in the *Yucatan* region of Mexico. They found a 180 km diameter ring like structure named *Chicxulub crater*. Using the $^{40}\mathrm{Ar}/^{39}\mathrm{Ar}$ method, the age of the crater was determined to be 65 million years.

Schuraytz et al. [220] found Ir anomalies even in subsplits of melt rock and melt breccia from the Chicxulub impact basin.

Several other craters appear to have been formed at about the time of the K-T boundary. As it has been observed during the Shoemaker-Levy 9 impact on Jupiter, an asteroid could fragment before its collision, therefore, the craters found were impacts of a larger body that fragmented before collision with the Earth. The craters (apart from the Chicxulub crater) that have been found are

1. Boltysh crater (24 km diameter, 65.17 ± 0.64 Ma old), Ukraine.
2. Silverpit crater (20 km diameter, 60–65 Ma old) in the North Sea.
3. Eagle Butte crater (10 km diameter, $<$65 Ma old) in Alberta, Canada.
4. Vista Alegre crater (9.5 km diameter, $<$65 Ma old) in Paraná State, Brazil.

5.3.2.2 Mass Extinction During the KT-Event

The biological system is quite complex and the extinction of one group inevitably leads to extinction of other groups.

The K-T impact caused a major change in both marine and land ecosystems. Before the K-T extinction, about 50% of known marine species were sessile, and after it, only about 33% were sessile. On land, the *dinosaurs* became extinct, therefore,

mammals were able to become the dominant land vertebrates; this seems important, also, for human evolution.

In North America, as many as 57% of plant species may have become extinct. The Paleocene recovery of plants began with a "fern spike" like that which signals the recovery from natural disasters (e.g., the 1980 Mount St. Helens eruption). The effects were quite different for different organisms. Some trends can be stated:

- Organisms that depended on photosynthesis became extinct or suffered heavy losses – from photosynthesizing plankton (e.g., coccolithophorids) to land plants. And so did organisms whose food chain depended on photosynthesizing organisms, e.g., tyrannosaurs (which ate vegetarian dinosaurs, which ate plants).
- Organisms which built calcium carbonate shells became extinct or suffered heavy losses (coccolithophorids; many groups of molluscs, including ammonites, rudists, freshwater snails, and mussels). And so did organisms whose food chain depended on these calcium carbonate shell builders. For example, it is thought that ammonites were the principal food of mosasaurs.
- Omnivores, insectivores, and carrion-eaters appear to have survived quite well. At the end of the Cretaceous, there seem to have been no purely vegetarian or carnivorius mammals. Many mammals, and the birds which survived the extinction, fed on insects, larvae, worms, snails, etc., which in turn fed on dead plant matter. So, they survived the collapse of plant-based food chains because they lived in "detritus-based" food chains.
- In stream communities few groups of animals became extinct. Stream communities tend to be less reliant on food from living plants and are more dependent on detritus that washes in from land. The stream communities may also have been buffered from extinction by their reliance on detritus-based food chains [225].
- Similar but more complex patterns have been found in the oceans. For example, animals living in the water column are almost entirely dependent on primary production from living phytoplankton. Many animals living on or in the ocean floor feed on detritus, or at least can switch to detritus feeding. Extinction was more severe among those animals living in the water column than among animals living on or in the sea floor. No land animal larger than a cat survived.
- The largest air-breathing survivors, crocodilians and champsosaurs, were semi-aquatic. Modern crocodilians can live as scavengers and can survive for as long as a year without a meal. And modern crocodilians' young are small, grow slowly, and feed largely on invertebrates for their first few years – so they rely on a detritus-based food chain.

It is not clear how long the K-T extinction took. Some theories require a rapid extinction (few years to a few 10^3 years), others require longer periods. It has also been argued that some dinosaurs survived into the Paleocene. This favors a gradual extinction of the dinosaurs. But this seems now very unlikely because all the remains found are fragments which could have been reworked.

Pope, D'Hondt, and Marshall [191] claimed that mass extinction of marine plankton appeared abrupt and right at the K/T boundary. Marshall and Ward (1996) found a major extinction of ammonites at or near the K-T boundary, a smaller and

slower extinction of ammonites associated with a marine regression shortly before that, gradual extinction of most inoceramid bivalves well before the K-T boundary, and a small, gradual reduction in ammonite diversity throughout the very late Cretaceous. This analysis may favor the idea that several processes contributed to the mass extinction in the late Cretaceous seas.

5.3.2.3 The Impact and its Consequences

The asteroid that has impacted near the coast must have caused gigantic tsunamis. Evidence for such a scenario has been found all round the coast of the Caribbean and eastern USA – marine sand in locations that were then inland, and on the other hand, vegetation debris and terrestrial rocks in marine sediments dated to the time of the impact.

The crater's shape suggests that the asteroid landed at an angle of 20° to 30° from horizontal travelling north-west. This would have directed most of the blast and solid debris into the central part of what is now the USA.

Some of the most severe consequences were as follows:

- Global dust cloud: this blocked sunlight and photosynthesis became reduced for years. The extinction of plants and phytoplankton as well as of all organisms dependent on them was a consequence. This includes also the predatory dinosaurs and herbivores. However, it is clear that organisms whose food chain were based on detritus could have survived. Moreover, the asteroid landed in a bed of gypsum (calcium sulphate), which would have produced a vast sulphur dioxide aerosol. This would have further reduced the sunlight.
- Global firestorms: after the impact, when fragments fell back to the Earth, global firestorms resulted. Fluid inclusion in amber suggest a higher oxygen content, thus, combustion would have been supported. The widespread fires would have increased the CO_2 content in the atmosphere (also because of lack of oxygen production). Therefore, after the long winter due to sunlight blocking, a hot atmosphere due to increased greenhouse gas concentration would have contributed to extinction of many species later than at impact time. This was studied, e.g., in the paper of Lyons et al. [151]. Their calculations also show that an object of 10 km size with impact velocity 20 km/s would produce ejecta mass of 2.0×10^{14} kg and cause a global layer of 0.17 mm thickness.[8] Their conclusion was that the re-entry of 1 mm sized ejecta did not ignite fires globally.
- Acid rain: this is now believed to have been of minor importance because animals that are vulnerable to acid rain (e.g., frogs) survived. Dust cloud and aerosols (sulphuric) would wash out of the atmosphere within 10 years.

[8] This does not include the mass of the impacting object.

5.3.3 Controversy about the K-T Impact Theory

There is still some controversy whether the K-T extinction event was really caused by an impact or whether this impact was the main cause for the mass extinction.

The Deccan Traps flood basalts (massive volcanic eruptions in India; they covered an area roughly the size of France) could have also caused mass extinction. It seems now established [106] that 2/3 of the Deccan Traps occurred over a time span of 1 million years about 65.5 Ma (million years ago) causing a rapid extinction over some 10^3 years. It should be also noted that a very large crater, 450–600 km, has been recently reported in the sea floor off the west coast of India, the so-called *Shiva crater*. The Shiva crater has been dated at about 65 Ma. It could have triggered the Deccan volcanoes. However, this feature has not yet been accepted by the geologic community as an impact crater and may just be a sinkhole depression caused by salt withdrawal.

Another important event at the K-T boundary is the so-called *Maastrichtian Sea-level Regression*. This can be found in Maastrichtian rock sequences: earlier rocks represent shorelines and even seabeds and latest rocks are terrestrial. This regression is not completely understood. One explanation could be that the mid-ocean ridges became less active and, therefore, sank under their own weight. A regression would have reduced the continental shelf area, which is the most species-rich part of the sea, and it would have caused climate changes, e.g., by disrupting winds and ocean currents and reducing the Earth's albedo. Marshall and Ward's analysis [156] suggests this was not enough to exterminate the ammonites. Keller [120] discusses the late Maastrichtian mantle plume volcanism and gives implications for impacts and mass extinctions.

The supernova hypothesis for the K-T extinction can be nearly excluded because ^{244}Pu should have been produced (half-life 81 Myr) but this isotope is not found in these layers.

It has been also suggested that mass extinction may arise through a purely biotic mechanism as the result of so-called coevolutionary avalanches (e.g., Newman [180]). Another ecological model for extintcion was given by Abramson [1].

5.3.4 The Permian-Triassic Event, P-T Event

Another mass extinction occurred at the Permian-Triassic boundary (PTB). This extinction can be also explained by an impact. The impact crater could have been the Australian Bedout crater. This crater is 250.7×10^6 years old, which coincides with the P-T-mass extinction that occurred 250 million years ago. There is strong evidence that this crater has been formed by an impact of an asteroid of about 10 km size because shocked quartz is found which cannot be of volcanic origin.[9] Moreover,

[9] Extreme volcanic activity fractures quartz only in one direction, but shocked quartz is fractured in several directions.

shocked quarz found at the Chicxulub crater shows very similar properties like the samples found here.

Becker et al. [11] report about extraterrestrial Cr isotopes as an evidence for the impact theory of the Bedout crater. Basu et al. [11] found multiple chondritic meteorite fragments in rock sample in Antarctica, dating to the end of the Permian. They explain this and other meteoritic fragments found in an end-Permian bed in southern China as indications of a global impact event at the Permian-Triassic boundary.

Poreda and Becker [192] investigated fullerenes with extraterrestrial noble gases in the PTB[10] at Graphite Peak, Antarctica and PTB fullerenes from Meishan, China, and Sasayama, Japan. They showed their similarity with zero-age-deep-sea interplanetary dust particles (IDPs) and some particle from the KT-boundary.

As for the K-T event, there are also indications of massive lava outflows in Siberia. These outflows could have been triggered or enforced by the impact of the asteroid.

5.3.5 Mass Extinctions by Flood Basalt Volcanism

Besides impacts, major mass extinctions in Earth's history are also attributed to flood basalt volcanism as it was pointed out already.

Flood basalts are generally linked to hot spots or superheated mantle plumes and they cover areas of 0.2×10^6 to 2.0×10^6 km^2 and the thickness can be up to 1 km.

Large igneous provinces are thought to be caused by the arrival of a mantle plume in the Earth's outermost layer, the lithosphere. The plumes are proposed richer in lighter elements and hotter than the surrounding mantle. As they rise, magma (liquid rock) is generated by partial melting of the plume material. The magma is injected into the lithosphere and erupted onto the Earth's surface to form huge basalt lava flows. The first few million years of a newly arrived mantle plume seem the most fertile in terms of magma production, and flood basalts are, therefore, formed in a very short period of geological time. The surface manifestations of mantle plumes are often called hotspots. Plumes are thought to originate very deep in the Earth – perhaps at the core-mantle boundary for the larger ones and at a depth of about 600 km for the smaller ones – but they are probably related to the breakup of continents (rifting), so there is some influence from global plate tectonic processes.

How can such catastrophic events be traced? One example is Deep Sea Drilling Project (DSDP) Site 216 on Ninetyeast Ridge in the Indian Ocean, which tracks the passage of the oceanic plate over a superheated mantle plume during the late Maastrichtian, resulting in lithospheric uplift, the formation of islands built to sea level, and volcanic activity lasting more than 1 million years. DSDP Site 216 was cored on the crest of Ninetyeast Ridge, which is currently located just north of the equator (lat. $1^o27.73'$ N, long. $90^o12.48'$) and at a water depth of 2,237 m. During the late Maastrichtian Site 216 was located at about 40°S (see Keller [119]). The

[10] Permian Triassic Boundary.

Maastrichtian is named after the Dutch city Maastricht, where many fossils from that period were found and it is the last stage of the Cretaceous period from 70.6 $\pm$ 0.6 Ma to 65.5 $\pm$ 0.3 Ma.

Rampino and Stothers [199] gave an overview over flood basalt volcanism during the past 250 million Years. The age of the basalt floods can be dated from K-Ar and Ar-Ar methods. In total, they found 11 distinct periods during the past 250 Myrs and they argue that the episodes came quasiperiodically with a mean cycle time of 32 million years.

A compilation of younger continental flood basalts is given in Table 5.3.

What would be the environmental effects of flood basalts eruptions? This question is extremely difficult to answer because there are several influences on climate acting in cooling or warming:

- Climatic cooling from sulfuric acid aerosols and sulfur volatiles (SO_2).
- CO_2 and SO_2 gases could cause a greenhouse warming.
- Acid rains.
- Changes in the ocean chemistry.
- Changes in the ocean circulation.
- Changes in the ocean's oxygenation.

The last three items become extremely important when basaltic volcanism becomes associated with large submarine plateaus, i.e., basalt eruption in an oceanic environment.

The appearance of the Ethiopian-Afar plume head at the Earth's surface, erupted approximately 30 Myr ago, over a period of 1 Myr or less. This was about the time of a change to a colder and drier global climate, a major continental ice-sheet advance in Antarctica, the largest Tertiary sea-level drop, and significant extinctions [105]. At that time, maxima of δ^{18}O value were measured. This also indicates a significant cooling and a minimum in the diversity of mammals. Injection of sulfur rich aerosols and dust in the atmosphere was also observed.

Table 5.3 Flood Basalt Provinces of the last 250 Myrs (adapted from http://www.geolsoc.org.uk)

Province	Age (Myr)	Volume (10^6 km^3)	Paleolatitude	Duration (Myr)
Columbia River	16 1	0.25	45°N	$\sim$ 1 (for 90%)
Ethiopia	31 $\pm$ 1	$\sim$1.0	10°N	$\sim$ 1
North Atlantic	57 $\pm$1	$>$1.0	65°N	$\sim$ 1
Deccan	66 $\pm$1	$>$2.0	20°S	$\sim$ 1
Madagascar	88 $\pm$ 1	?	45°S	$\sim$ 6 ?
Rajmahal	116 $\pm$ 1	?	50°S	$\sim$2
Serra Geral/ Etendeka	132 $\pm$ 1	$>$1.0	40°S	$\sim$ 1 or $\sim$5?
Antarctica	176 $\pm$ 1	$>$0.5	50-6°S	$\sim$ 1?
Karoo	183 $\pm$ 1	$>$2.0	45°S	0.5–1
Newark	201 $\pm$ 1	$>$1.0?	30°N	$\sim$ 0.6
Siberian	249 $\pm$1	$>$2.0	44°N?	$\sim$ 1

A simulation with modern climate models was carried out to study the effects of the Toba eruption, which was 75,000 years ago. For that simulation, a volcanic eruption 100 times larger than the 1991 Pinatubo eruption was assumed. The result was a global cooling for several years of up to 10°C but the aerosols left the atmosphere quickly and the climate recovered in less than a decade. Thus, even such strong single eruptions may not explain mass extinctions [211]. It is assumed that this eruption possibly wiped 70–99% of humans but on the other hand it helped Moderns to spread from Africa.

Glickson [75] suggest the dominant role of continental flood basalts and explosive CO_2 rich volcanic pipes for mass extinctions. He studied three overlaps between the ages of extraterrestrial impact and volcanic events and arrived at the conclusion that both catastrophic events in combination could have caused mass extinction and large impacts on a thin oceanic crust could have triggered regional to global volcanic events.

5.4 NEOs

5.4.1 Classification and Definition

We now consider objects that could approach to small distances to the Earth and become possible impactors.

Chapman and Morrison [36] stated for the first time that although impacts on Earth by asteroids and comets are extremely infrequent as seen from personal experience, the long-term statistical hazard is comparable to that of many other more familiar disasters. They also raised the question of whether mitigation measures should be considered.

The near-Earth asteroid 433 Eros (approximately $33 \times 13 \times 13$ km size) had been discovered in 1898, and in the 1930s, similar objects have been found. In order of discovery, these were as follows: 1221 Amor, 1862 Apollo, 2101 Adonis, and finally 69230 Hermes, which approached Earth in 1937 (0.005 AU). At that time, astronomers began to realize the possibilities of Earth impact.

The US military also declassified the information that its military satellites, built to detect nuclear explosions, had detected hundreds of upper-atmosphere impacts by objects ranging from 1 to 10 m across.

The term NEOs denotes Near Earth Objects. These include comets and asteroids that have changed their orbits due to gravitational interactions with other planets. *Comets* originally formed in the cold outer solar system, they are mainly composed of water ice and rocks, *asteroids* mainly formed in the inner solar system.

NEOs are further subdivided into the following groups (q denotes the perihelion distance):

- NECs: Near-Earth Comets, $q < 1.3$ AU, P < 200 years.
- NEAs: Near-Earth Asteroids, $q < 1.3$ AU.

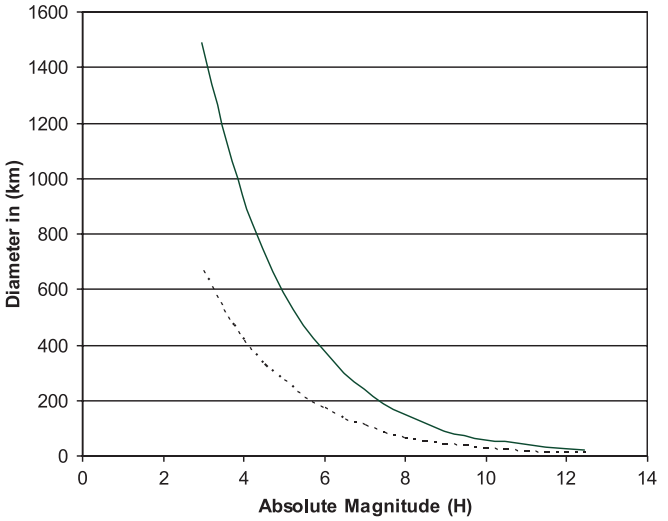

Fig. 5.8 Relation between absolute magnitude of an asteroid and its diameter for albedo 0.25 (bright object, *dashed line*) and albedo 0.05 (dark onject, *full line*). An object at a distance of 1 AU to the Earth and the Sun would be seen with the naked eye if it has a diameter of more than 600 km (low albedo) or 200 km (high albedo)

- Atens: Earth-crossing NEAs with semi-major axes smaller than Earth's (named after asteroid *2062 Aten*). $a < 1.0\,\mathrm{AU}, \mathrm{q} > 0.983\,\mathrm{AU}$
- Apollos: Earth-crossing NEAs with semi-major axes larger than Earth's (named after asteroid *1862 Apollo*). $a > 1.0\,\mathrm{AU}, \mathrm{q} < 1.017\,\mathrm{AU}$
- Amors: Earth-approaching NEAs with orbits exterior to Earth's but interior to Mars' (named after asteroid *1221 Amor*). $a > 1.0\,\mathrm{AU}, 1.017 < \mathrm{q} < 1.3\,\mathrm{AU}$
- PHAs (Potentially Hazardous Asteriods): NEAs whose Minimum Orbit Intersection Distance (*MOID*) with the Earth is 0.05 AU or less and whose absolute magnitude (*H*) is 22.0 or brighter. $\mathrm{MOID} \leq 0.05\,\mathrm{AU}, \mathrm{H} \leq 22.0$

Currently, about 800 PHAs are known.

The absolute magnitude is the visual magnitude an observer would record if the asteroid were placed at a distance of 1 AU from the Sun and the Earth. In Fig. 5.8, the relation between the absolute magnitude of an asteroid and its diameter is given for albedo 0.25 (bright object, dashed curve) and 0.05 (dark object, full line). The smaller the albedo, the darker the object. For comparison: the Moon has a very low albedo (0.07), while Venus has a high albedo (0.60) because of the high reflectance of its clouds.

5.4.2 Orbit Instabilities

In the above section, we have defined several families of orbits of comets and asteroids. There are cases for which no real distinction between comets and asteroids

can be made. In this section, we will discuss how stable the orbits of these object families are.

Levison and Duncan [141] discuss how Transneptunian Objects (TNO) can penetrate to the region of the Centaurs by gravitational interaction with the giant planets. They show that the Jupiter family comets can be transformed into long lived NEOs. About 100 objects are known to belong to the *Centaur population* (perihelia and semi-major axes between the orbits of Jupiter and Neptune). Kovalenko, Churyumov, and Babenko [127] investigated whether these objects could be a source of NEOs. They made numerical calculations for a time interval $\pm 10^6$ years for the objects. The main results were that objects with perihelia $< 5\,\mathrm{AU}$ evolved cometary like activity and 10% of the orbits in the past and 15% of the orbits in the future reach perihelia $< 5\,\mathrm{AU}$. Thus, it is possible that they contribute to the NEOs.

Most NEOs are supplied from the main asteroid belt. Long term integrations of hypothetical *Jupiter Family comets* show that these objects could contribute to the NEO population. Ninety percent of decoupled NEOs could have their origin in the main asteroid belt (see Harris [94]). NEOs could be fragments arising from collisions in the main asteroid belt. Menichella, Paolicchi, and Farinella [164] estimate that about 370 km sized objects and 2.5×10^5 0.1 km sized objects can be injected per million years from the main belt to the typical orbits of NEOs. This gives an estimate of the steady state population of NEOs. Morrison [171] and Rabinowitz et al. [196] give the following values: from the main belt asteroids about 2,000 km sized objects and 10^6 objects $> 1\,\mathrm{km}$. Thus, the asteroid belt is a major source for NEOs.

When a NEO comes close to the Earth, a collision may occur. Another possibility is a capture by the Earth-Moon system [45].

5.5 Impact Risk Scale

5.5.1 Surveillance Systems

There are several systems of robotic telescopes and observatories that automatically survey the sky for asteroids. Great progress of such systems was made by introducing large CCD cameras and with the appropriate software many detections are possible during one night.

As an example, we see in Table 5.4 the number of discovered objects for different dates.

Some of these monitoring systems are listed below:

- The Lincoln Near-Earth Asteroid Research (LINEAR[11]) team. The goal of the LINEAR program is to demonstrate the application of technology, originally developed for the surveillance of earth orbiting satellites, to the problem of detecting and cataloging NEAs. By the year 2004, more than 12 million observations

[11] http://www.ll.mit.edu/LINEAR/

Table 5.4 Detection of NEOs and other objects (http://neo.jpl.nasa.gov/stats/). NEC are near earth comets, PHA-km denotes PHAs with approximately 1 km diameter

Date	NEC	Aten	Apollo	Amor	PHA km	PHA	NEA km	NEA	NEO
2007									
08-22	64	390	2,491	1,822	136	866	720	4,711	4,775
2006									
08-01	61	328	2,165	1,603	130	780	685	4,102	4,163
2005									
08-01	59	283	1,802	1,347	126	693	646	3,436	3,495
2004									
08-01	55	226	1,500	1,156	117	598	601	2,885	2,940

have been reported to the MPC (Minor Planet Center). Currently, LINEAR telescopes observe each patch of sky 5 times in one evening with most of the efforts going into searching along the ecliptic plane where most NEOs would be expected.

- The Near-Earth Asteroid Tracking (NEAT) team, located in White Sands, USA; they use a pair of 1 m telescopes.
- Spacewatch, carried out by the Lunar and Planetary Laboratory, University of Arizona, USA. The two telescopes used have diameters of 1.8 m and 0.9 m.
- The Lowell Observatory Near-Earth-Object Search (LONEOS) uses a 0.6 m f/1.8 Schmidt Telescope equipped with a 4 k $\times$ 4 k CCD. The detector, thus, covers a huge field of 2.9 $\times$ 2.9 degrees. Each month four scans over the entire sky are made.
- The Catalina Sky Survey (CSS) in the Catalina Mountains, north of Tucson equipped with two telescopes 68/76 cm f/1.9 (classical Schmidt); also, the Mt. Lemmon Sky Survey (1.5 m telescope, performs spectrophotometry to determine the chemical composition of the object) and the Siding Spring Survey (NSW Australia, 0.5 m Schmidt) are included there.
- The Campo Imperatore Near-Earth Objects Survey (CINEOS) (Schmidt 90 cm, f/1.8, operated by ROSI (Rome Observatory Imager)).
- The Japanese Spaceguard Association, it operates the Bisei Spaceguard Center and besides NEOs, space debris is also surveyed.
- The Asiago-DLR Asteroid Survey (ADAS), 67 cm Schmidt telescope.

The LINEAR system alone has discovered 71,770 asteroids.[12] The automated systems discovered 4,286 NEAs including over 600 that are more than 1 km in diameter (Fig. 5.9).

[12] Nov. 9, 2006.

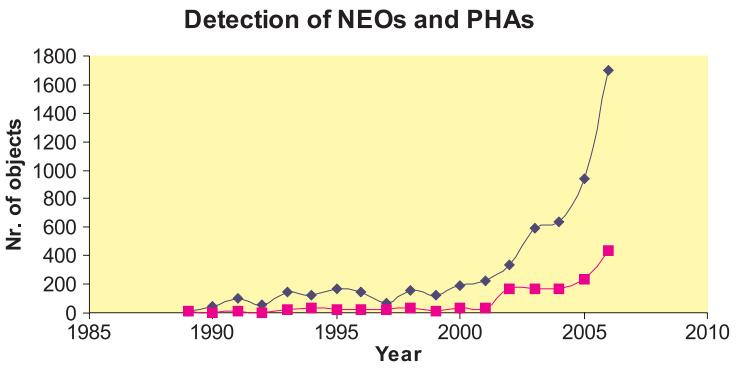

Fig. 5.9 Detection of NEOs (upper curve) and PHAs (lower curve) by Spacewatch

5.5.2 Torino Impact Scale

This scale was designed for public communication of impact risk, [19] and it is an outcome of the International Monitoring Programs for Asteroid and Comet Threat (IMPACT), a workshop held in Torino in 1999.

- No Hazard, white zone, **0**: The likelihood of a collision is zero, or is so low as to be effectively zero. It also applies to small objects such as meteors and bodies that burn up in the atmosphere as well as infrequent meteorite falls that rarely cause damage.
- Normal (Green zone), **1**: a routine discovery in which a pass near the Earth is predicted that poses no unusual level of danger. Current calculations show the chance of collision is extremely unlikely with no cause for public attention or public concern. New telescopic observations are very likely will lead to re-assignment to Level 0.
- Meriting Attention by Astronomers

  2: A discovery, which may become routine with expanded searches, of an object making a somewhat close but not highly unusual pass near the Earth. While meriting attention by astronomers, there is no cause for public attention or public concern as an actual collision is very unlikely. New telescopic observations very likely will lead to re-assignment to Level 0.
  3: A close encounter, meriting attention by astronomers. Current calculations give a 1% or greater chance of collision capable of localized destruction. Most likely, new telescopic observations will lead to re-assignment to Level 0. Attention by public and by public officials is merited if the encounter is less than a decade away.
  4: A close encounter, meriting attention by astronomers. Current calculations give a 1% or greater chance of collision capable of regional devastation. Most likely, new telescopic observations will lead to re-assignment to Level 0. Attention

by public and by public officials is merited if the encounter is less than a decade away.

- Threatening (Orange Zone):

  5: A close encounter posing a serious, but still uncertain threat of regional devastation. Critical attention by astronomers is needed to determine conclusively whether or not a collision will occur. If the encounter is less than a decade away, governmental contingency planning may be warranted.
  6: A close encounter by a large object posing a serious but still uncertain threat of a global catastrophe. Critical attention by astronomers is needed to determine conclusively whether or not a collision will occur. If the encounter is less than three decades away, governmental contingency planning may be warranted.
  7: A very close encounter by a large object, which if occurring in this century, poses an unprecedented but still uncertain threat of a global catastrophe. For such a threat in this century, international contingency planning is warranted, especially to determine urgently and conclusively whether or not a collision will occur.

- Certain Collisions (Red Zone):

  8: A collision is certain, capable of causing localized destruction for an impact over land or possibly a tsunami if close offshore. Such events occur on average between once per 50 years and once per several 1,000 years.
  9: A collision is certain, capable of causing unprecedented regional devastation for a land impact or the threat of a major tsunami for an ocean impact. Such events occur on an average between once per 10,000 years and once per 100,000 years.
  10: A collision is certain, capable of causing global climatic catastrophe that may threaten the future of civilization as we know it, whether impacting land or ocean. Such events occur on average once per 100,000 years or less often.

5.5.3 Palermo Technical Impact Hazard Scale

The Palermo Technical Impact Hazard scale (see Chesley et al. [38] combines two types of data – probability of impact and estimated kinetic yield – into a single "hazard" value.

The Palermo Scale value P is defined as the ratio of the impact probability P_i to the background impact probability over time in years T to the event:

$$P = \log_{10} \frac{P_i}{f_B T} \tag{5.3}$$

A rating of 0 means the hazard is as likely as the background hazard (defined as the average risk posed by objects of the same size or larger over the years until the date

of the potential impact). A rating of +2 would indicate the hazard is 100 times more likely than a random background event.

The annual background impact frequency is defined for this purpose as

$$f_B = 0.03E^{-0.8} \tag{5.4}$$

in which the energy threshold E is measured in megatons.

The record for Palermo scale values (up to the end of 2006) is held by asteroid (29075) 1950 DA, with a value of 0.17 for a possible collision in the year 2880.

The topic of NEO Impact Hazard Scales in the Context of Other Hazard Scales was discussed by Chapman and Mulligan [37].

5.6 Collisions and Habitability

5.6.1 Impacts of Comets

Comets are mainly made of ice, whereas asteroids are made of rock and no ice (some of metals). There is a large uncertainty in the prediction whether a comet would collide with a planet, say, the Earth or not because

- For newly discovered comets, it is difficult to determine precisely their orbital elements because there are strong residuals from the position of the background stars.
- As the comet approaches the Sun, its surface warms up and gases (like geysers) burst out of the surface. This adds so-called *nongravitative forces*, which are unpredictable.

It can be estimated that at best, a reliable prediction can be made only 20 days in advance whether a comet will really collide with Earth or not.

In Fig. 5.10, comet *Temple1* is shown as seen from a NASA spacecraft flyby on 4 July 2005. On that day, the comet was hit by a 370 kg projectile.

5.6.1.1 Shoemaker-Levy comet impact on Jupiter

Comet Shoemaker-Levy was detected by Eugene and Carolyn *Shoemaker* and David *Levy* on a photograph on 24 March 1993.[13] It was then found that the comet had a very close approach with Jupiter on 7 July 1992 (the distance was only $1.4\,R_J$). Due to the strong tidal forces, the comet disrupted into several fragments.

In July 1994, we had the unique opportunity to witness a cometary impact on Jupiter. The comet *Shoemaker-Levy 9* fragmented and it could be easily observed

[13] The photograph was made with a small 0.4 m telescope.

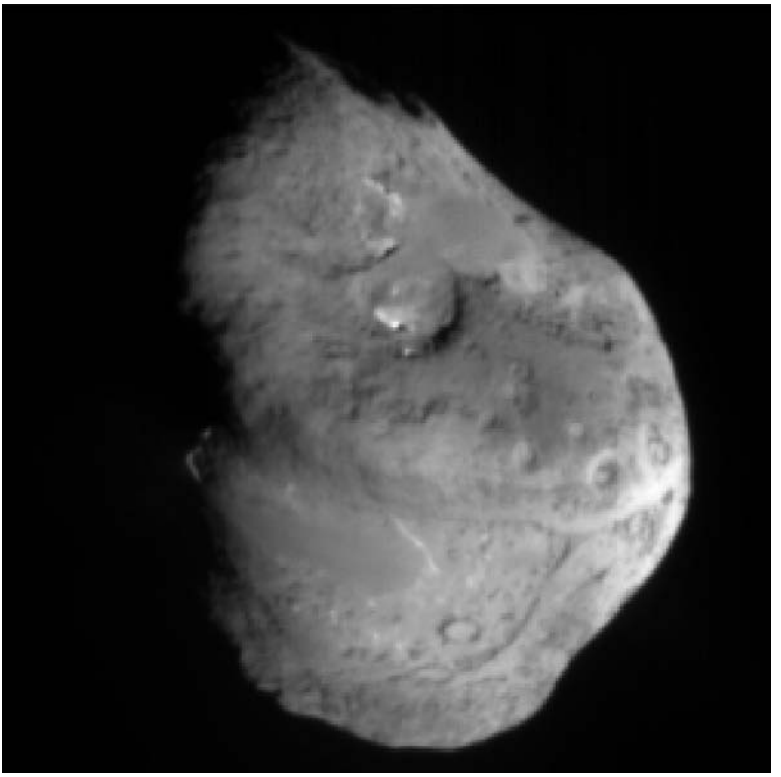

Fig. 5.10 Comet Temple 1. This image shows comet Temple 1 approximately 5 minutes before Deep Impact's probe smashed into its surface. It was taken by the probe's impactor targeting sensor. The Sun is to the right of the image. Smooth regions with no features (lower left and upper right) are probably younger than rougher areas with circular features, which are most likely impact craters. The nucleus is estimated about 5×7 km . Image credit: NASA/JPL-Caltech/UMD

even with amateur telescopes how these impacts changed the atmosphere of Jupiter (Fig. 5.11). From July 16 through July 22, pieces of the comet collided with Jupiter. More than 20 fragments were discernable with diameters up to 2 km.

The energy of the impacting fragments was estimated to about 10^7 megatons TNT.[14]

$$1\,\mathrm{kT\ TNT} = 4.184 \times 10^{12}\,\mathrm{J} \tag{5.5}$$

The sum of all conventional bombs that were delivered in the Second World War had an equivalent of 2 MT TNT. Thus, the energy released by only one fragment of the Shoemaker-Levy Jupiter impact was 10^7 times above this value. It was even possible to observe this collision from the Galileo spacecraft which observed fireballs up to 80 s. Near the beginning, the temperature was near 3,000 K and after

[14] The Hiroshima bomb released 15 kt

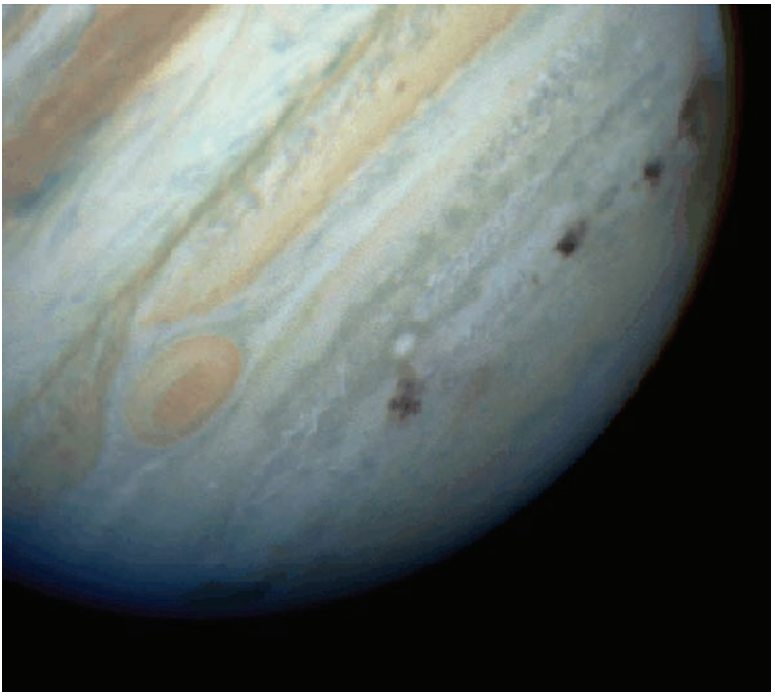

Fig. 5.11 Comet Impacts on Jupiter Picture Credit: NASA, Hubble Space Telescope

1 minute it dropped to 1,000 K. The so-called G-fragment, e.g., released an energy of 2.5×10^{19} J [34]. The released energy values depend on the model used for the comet (e.g., if it consists only of ice). It can also be estimated that such impacts occur on Jupiter every $\sim 10^3$ years. Thus, they are about 10^5 times more frequent than the Cretaceous-Tertiary extinction energy impactor on Earth ($\sim 10^8$ yr) [242].

Moreno, Marten, and Matthews [169] were able to detect molecular traces of that impact in Jupiter's stratosphere (HCN, CO, CS).

Comets are often split under the strong gravitational forces by the planets. Another example is comet 73P/Schwassmann-Wachmann 3. Its nucleus split into four fragments in December 1995. Shoemaker and Izett [227] claimed that from the stratigraphy of the *K/T boundary*, there is evidence for multiple impacts and a possible comet stream. This could be expected remembering the Shoemaker-Levy comet impact on Jupiter was that the comet broke into several fragments before colliding with the planet.

There is a difference between comets that split because of tidal effects and those that split due to nongravitational forces (outgassing). For nontidally split comets, the principal nucleus has a leading position and the other fragments trail behind it [222].

5.6.2 Probability of Cometary Impacts

What are the differences between impacts from comets and asteroids (see also Weissman [263])?

- Comets have higher approach velocities.
- The mean impact velocity for comets is about $51.8\,\mathrm{km\,s^{-1}}$.
- Terrestrial encounters with long period comets are difficult to predict. They enter the inner planetary system unpredictably through random perturbations in the Oort cloud.
- Warning times for cometary impacts are short, maybe only a few months.
- The mean impact probability for Earth-crossing long-period comets is 2.2×10^{-9} per perihelion passage, assuming a uniform perihelion distribution and random inclination distribution for comets interior to 1 AU.

For *Jupiter-family comets*, whose returns are predictable (once discovered), only 22 Earth-crossers are known (excluding the many fragments of 73P/Schwassmann-Wachmann 3). Of these, four are lost, eight have only been observed on one return, and one is no longer Earth-crossing. Their mean impact probability is 7.3×10^{-9} per orbit or 1.3×10^{-9} per year and their mean encounter velocity with the Earth is $22.9\,\mathrm{km\,s^{-1}}$, with a most probable encounter velocity of $19.9\,\mathrm{km\,s^{-1}}$.

For *Halley-type comets*, whose returns are also predictable, another 16 Earth-crossers are known, of which one is lost and six have not yet made a second observed appearance. Their mean impact probability is 7.0×10^{-9} per orbit but only 0.16×10^{-9} per year because of their longer orbital periods. Their mean encounter velocity is $45.4\,\mathrm{km\,s^{-1}}$, with a most probable encounter velocity of $52.3\,\mathrm{km\,s^{-1}}$. For further references see Bottke and Morbidelli [23], who gave an overview (over the last 4 Gyears) on the asteroid and comet impact flux in the terrestrial planet region.

Rickman et al. [208] estimated the cometary contribution to planetary impact rates. The authors find a terrestrial impact rate of about 1×10^{-6} per year. This is already significant ($\sim 20 - 50\%$) compared with the total estimated impact rate by km-sized bodies. The value might be even higher when contributions from Halley-type and other long-period comets are considered. The conclusion, therefore, is that comets yield a large, perhaps dominant, contribution to km-sized terrestrial impactors.

Levison et al. [142] also investigated planetary impact rates from ecliptic comets. Their simulation led to the result that 21% of the objects that impact *Jupiter* are bound to the planet before impact. For Saturn, the fraction of bound impactors is much lower.

Terrestrial impact probabilities for parabolic comets were treated in the paper of Steel [234]. He assumed a theoretical distribution of parabolic comets arriving from a spherical source. The mean impact probability was found to be 2.2×10^{-9} per perihelion passage. The mean impact speed was 55 km/s. Observations of long-period comets suggest a distribution that varies $\sim \sqrt{q}$, which has only tiny effects on the numbers: the mean impact probability in that case would be 2.6×10^{-9} per perihelion passage and the mean impact speed would be slightly higher.

Roddy, Shoemaker, and Anderson [212] estimated the energy of formation and the possible size of the impacting object that led to the formation of the Manson crater (located near Manson, Illinois; impact occurred 75 million years BP).

Ong, Asphaug, and Plesko [183] discuss how water was delivered to the Moon by comet impacts.

A variety of mechanisms have been proposed to explain why comets get perturbed into highly elliptical orbits: close approaches to other stars as the Sun follows its orbit through the Galaxy; A Sun's hypothetical companion star Nemesis; or an unknown Planet X.

5.6.3 Impacts and Their Consequences

It is evident that impacts of comets or asteroids or even large meteorites affect the habitability of a planet. Geological and astronomical evidence for comet impact and comet showers during the last 100 million years were reviewed by Shoemaker [226].

5.6.3.1 Effects in the Atmosphere and Magnetosphere

The ionospheric and magnetospheric disturbances caused by impacts of small comets and asteroids on Earth were investigated by Kovalev, Nemchinov, and Shuvalov [128]. They made two-dimensional hydrodynamical computations of the disturbances caused by the impacts of objects with diameters from 50 m to 2 km as they enter and pass the atmosphere. During the passage, an upward plume is formed by heated air and ablation products. This causes magnetic variations comparable to a magnetic storm for objects similar to the Tunguska meteorite and for a 1-km body, the energy of the high-velocity plume is comparable to that of the Earth's magnetic field. Extremely intense magnetospheric disturbances are a consequence. We have already mentioned what a weakening of the Earth's magnetic field means for habitability: the protection against charged high-energetic particles is destroyed and ionizing particles produce mutations in cells.

5.6.3.2 Effects on the Surface

Endoliths are organisms (they are found among certain Bacteria, Archaea, and even Fungi) that live inside rocks or in pores between mineral grains (Fig. 5.12). By asteroid and comet impacts, the habitats available can be changed. Cockell et al. [42] describe how such impacts lead to rock porosity, fracturing that can result in the formation of cryptoendolithic habitats. Chasmoendolithic habitats are associated with cracks and cavities connected to the surface of the rock and are commonly increased in abundance as a result of impact bulking.

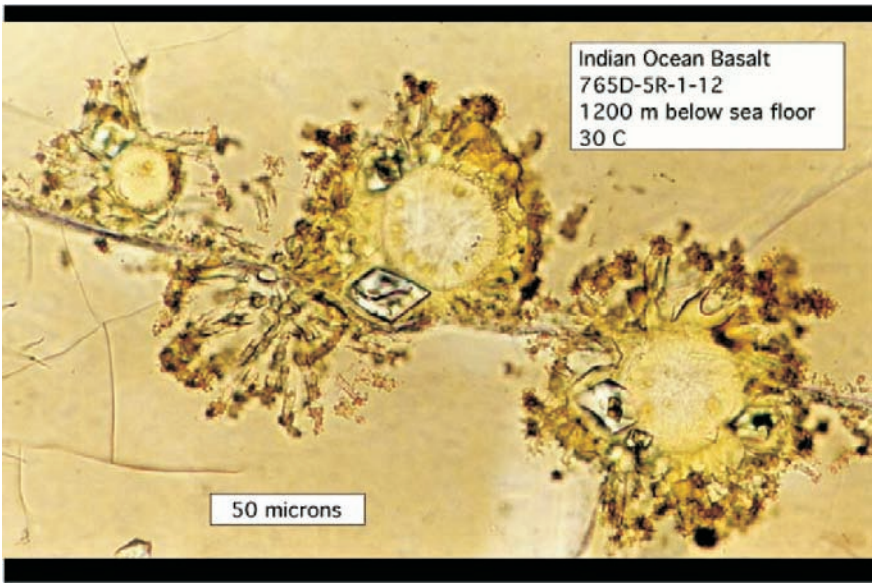

Fig. 5.12 Rock-eating endolith, taken from a basalt found 1,200 m beneath the Indian Ocean floor. Image courtesy of NOAA Ocean Explorer

For our civilization, it is of course most interesting what would happen if a comet or asteroid impact on our populated earth. Computer models have been made by Lewis [143]. Yeomans [271] also described such a scenario. Besides of the catastrophe caused by such an impact, it is stressed that these impacts may have provided the building blocks for the formation of life as has been mentioned before. The mass extinction 65 million years ago gave the mammals the predominant position on Earth. It is also mentioned that in the next century, comets and asteroids could supply many raw materials.

Rampino and Stothers [200] gave a review on mass extinctions and cometary impacts.

Tsunami from Asteroid-Comet impacts were studied by Hills et al. [104]. These authors investigated the effect of an asteroid impact of 5–6 km diameter in the middle of the large ocean basins. We cite their main results here: areas with gradual continental shelves such as the British Isles and Florida were found less susceptible to impact tsunami, while regions with little continental shelf are more exposed. An asteroid 5–6 km in diameter hitting in mid-Atlantic would produce tsunami that would travel all the way to the Appalachian Mountains in the upper two-thirds of the United States; e.g., Washington, DC would be washed out. In Florida, the damage would be much less except in the area from West Palm Beach through Ft. Lauderdale to Palm Beach where a deep-water shipping channel would funnel the tsunami energy to shore. In Europe, the most susceptible area is the Spain-Portugal peninsula, which has very little continental shelf. In retrospect, this result is not surprising as this peninsula has been subject to earthquake tsunami during

the last few hundred years. An asteroid 5–6 km in diameter impacting between Hawaii and the West Coast would wash out the coastal cities. The area between Los Angeles and San Diego is particularly vulnerable. This same impact would produce tsunami that would cover most of the land area of Oahu, the most populous Hawaiian island.

5.6.3.3 Effects on Other Terrestrial Planets

Owen and Bar-Nun [185] proposed a model for the delivery of volatiles to the inner planets by comets. The observed volatile inventory on *Venus*, *Earth*, and *Mars* can be explained by cometary impacts and also several epochs of clement conditions on early Mars can be deduced from these assumptions.

Grinspoon and Bullock [85] considered cometary impacts on *Venus*. They discuss the impact of gases like H_2O or SO_2 on the climate on Venus. In outgassing events, these gases can be released to the atmosphere. This contributes to a thicker atmosphere and increased greenhouse warming. The maximum warming occurs when SO_2 has been removed by reaction with surface minerals and H_2O is still present. Thus, warming takes place several hundred million years after the outgassing. They estimate that the atmosphere of Venus contains currently about 10^{15} kg of water. This amount corresponds to a comet of 25 km diameter. It is believed that comets could have been an important contributor to the current water content in the atmosphere of Venus.

Menchinov et al. [178] have shown that for velocities of the comet greater or about 20 km/s, a portion of energy emitted from the fireball exceeds 20% of the total energy of the explosion and this quantity does not change very much with the velocity.

5.6.4 The Tunguska Event

In June 1908, a massive explosion occurred near the Tunguska River in Siberia (Fig. 5.13). Passengers in the Transiberean Railway at a distance of about 500 km reported to have seen a huge fireball. Due to the isolation of this region and due to chaotic years (Russian Revolution, World War I), the first expedition to examine this event was made in 1921. In 1927, a big expedition started but did not find any traces of meteoritic iron nor a crater. However, a region of storched trees over more than $2,000\,\text{km}^2$ was detected. Later, this area was photographed but despite the devastation, no crater signs were seen. In the 1960s, expeditions found traces of microscopic glass spheres containing high proportions of nickel and iridium, which is typical for meteorites.

This lead to the widely accepted theory that a meteoroid exploded at a height between 6 and 10 km above the Earth's surface.

Fig. 5.13 Map of Russia with the location of the Tunguska event. Wikimedia Cosmos, GNU FDL

Let us determine the energy released. It is estimated that a stony meteoroid of about 10 m in diameter can produce an explosion of around 20 kt.[15] Explosions at the 20 kt range may occur yearly in the atmosphere. The Tunguska event can be compared with the largest H-bomb ever tested ~ 10 Mt. The frequency of such events is estimated once every 300 years.

What was the nature of the impacting body? In 1978, Kresak [129] suggested that the body was a piece of the short-period comet Encke. This hypothesis was criticized in 1983 by Sekanina [221]. His argument was that comets would completely disintegrate travelling through the atmosphere. However, the impacting Tunguska body was intact before disintegration. Farinella et al. [57] calculated that the object originated from the asteroid belt. The mystery remains, however, as to why there is no crater when the object was a stony asteroid. In any case, it is clear that the object must have had a diameter of about 80 m.

In June 2007, Lake Cheko was discovered, which is located only 8 km near the epicenter. This lake is not found in maps older than 1908.

Other more speculative theories about the Tunguska event are as follows: a black hole that passed through the Earth, a small piece of antimatter, a deuterium rich comet where fusion set in during his passage and heating through the atmosphere, and, of course, a crash of an UFO.

[15] This is comparable to the atomic bomb that destroyed Hiroshima.

Chapter 6
Catastrophes in Extrasolar Planetary Systems?

6.1 Collisions in Extrasolar Planetary Systems

After having discussed collisions in the solar system, we investigate whether such catastrophic events can be observed in extrasolar planetary systems. As we have shown, in the solar system, the most catastrophic events occurred during the formation phase when, e.g., the Moon was created by a collision of early Earth with a Mars-sized protoplanet. We then can raise several questions:

- Are such catastrophic collisions in the early phase of planetary system formation rare events?
- What amount of energy is released during such a collision?
- Is there any chance to observe such an event in an extrasolar planetary system during its formation stage?
- Are there certain types of planetary systems where catastrophic collisions occur more often than in other types?

The importance of such collisions is obvious because large satellites may be formed which stabilize the planet's rotation axis and because of the collision itself, the rotation speed of the object slows down.

6.1.1 Stability Catalogues

It is highly probable that extrasolar planetary systems contain at least one Jupiter-like object. As we have discussed, this is important because such a gas giant planet deflects comets that penetrate into the inner system where Earth sized planets in HZ might be located. Theoretical calculations were made for hypothetical planetary system containing such a Jupiter-sized object and terrestrial planets. Stability maps showed where Earth sized planets could be found at stable orbits. Sandor et al. [217] examined 15 known extrasolar planetary systems. HZ in extrasolar planetary systems were also investigated by von Bloh, Bounama, and Franck [253]. In that study, 86 extrasolar planetary systems were investigated and 60% of them could harbour habitable earth-like planets on stable orbits. In 18 extrasolar systems, the

A. Hanslmeier, *Habitability and Cosmic Catastrophes*, Advances in Astrobiology and Biogeophysics, DOI 10.1007/978-3-540-76945-3_6,

conditions even seem to be better than in our solar system. The orbital stabilities of the systems 51 Peg, 47 UMa, and HD 210277 were analyzed by Noble, Musielak, and Cuntz [181] and no stable orbits are found in the HZ of HD 210277.

Jupiter-sized planets might also be surrounded by habitable moons, as is assumed in our solar system (e.g. see the chapters about Europa and Titan).

6.1.2 Collision and X-ray Flashes

In this chapter, we investigate a catastrophe scenario: an Earth-sized object collides with a Jupiter-sized planet. The calculations show that two effects are to be expected:

- an immediate EUV to soft X-ray flash; this flash will last for several hours, thus, the detection probability is extremely low;
- long lasting IR afterglow, which could last for thousands of years.

Zhang and Sigurdsson [275] studied how such signals could be observed with existing and future technological facilities. It is shown that some of these events (both collisional flashes and post collision remnants) are really observable. Let us follow this paper in order to estimate some energies. A Jovian-like planet with mass $M_1 = 10^{30}, \mathrm{g} \sim \mathrm{M_J}$ and radius $R_1 = 7.0 \times 10^9, \mathrm{cm} \sim \mathrm{R_J}$ is hit by a terrestrial planet with mass $M_2 = 6 \times 10^{27}\,\mathrm{g} \sim \mathrm{M_E}$ and radius $R_2 = 6.4 \times 10^8\,\mathrm{cm} \sim \mathrm{R_E}$. M_J and R_J are and the mass and radius in units of the Jupiter values; M_E and R_E are mass and radius in units of Earth values.

The probability for glacing is

$$\sim R_2/R_1, \tag{6.1}$$

when the system is gravitational focused, and

$$\sim (R_2/R_1)^2 \tag{6.2}$$

when the colliding velocity largely exceeds the escaping velocity of the primary. For such a system, we obtain a probability for glacing between $\sim 0.09\frac{\mathrm{R_1}}{\mathrm{R_J}}$ and $0.01\left(\frac{\mathrm{R_1}}{\mathrm{R_J}}\right)^2$. Because of this low probability, we assume head-on collisions. Assuming zero velocity at infinity, the total energy of collision is given by

$$E_{\mathrm{coll}} = G\frac{M_2 M_1}{R_1} \sim 6 \times 10^{40}\,\mathrm{ergs}(\mathrm{M_E M_J/R_J}) \tag{6.3}$$

The relative velocity of impact becomes

$$v = \sqrt{2GM_1/R_1} \sim 44\,\mathrm{km/s}(\mathrm{M_J/R_J})^{1/2} \tag{6.4}$$

The interior sound speed of the impactors M_2 is

$$C_s = 10c_s(\,\mathrm{km/s}) \tag{6.5}$$

and $c_s \sim 1$ for terrestrial planets. Therefore, we have the classical shock situation:

$$v >> C_S \tag{6.6}$$

These shocks propagate both through M_1 and M_2 and most of the kinetic energy is converted to thermal energy. We can also estimate the shock crossing time for the impactor:

$$\tau_{\rm sh} \sim \frac{R_2}{C_s} \sim 640\,{\rm sec}\,({\rm R_E/c_s}) \tag{6.7}$$

This is equivalent to the rising timescale of a collision induced electromagnetic flash. One can also estimate the energy deposit rate:

$$\dot{E} \sim E_{\rm coll}/\tau_{\rm sh} \sim 9 \times 10^{37}\,{\rm ergs^{-1}}({\rm M_E M_J/r_J})({\rm R_E/c_s}) \tag{6.8}$$

The *Eddington luminosity* is the largest luminosity that can pass through a layer of gas in hydrostatic equilibrium. It can be derived easily. In case of hydrostatic equilibrium:

$$\frac{{\rm d}P}{{\rm d}r} = -g\rho = -G\frac{M\rho}{R^2} \tag{6.9}$$

The outward force of radiation pressure is

$$\frac{{\rm d}P}{{\rm d}r} = -\frac{\kappa\rho}{c}F_{\rm rad} = -\frac{\sigma_T\rho}{m_p c}\frac{L}{4\pi r^2} \tag{6.10}$$

σ_T is the Thomson scattering cross-section (for the electron; the gas is assumed purely ionized hydrogen). By equating (6.9) and (6.10), the *Eddington Luminosity* is

$$L_{\rm Edd} = 4\pi G M_P c/\sigma T \tag{6.11}$$

The Eddington limit for our planet is $L_{\rm Edd} \sim 5 \times 10^{34}$ ergs. Therefore, the energy deposit is a factor 10^3 above the Eddington limit. Half of the collisional energy is converted into internal energy (which is not radiated promptly). The other half is kinetic energy, a radiation supported expanding envelope is driven and oscillations and convection are induced in the planet's interior. This kinetic energy is thermalized later and radiated away over a long period (*IR afterglow*). The peak bolometric luminosity, however, becomes

$$L_{\rm peak} = \eta L_{\rm Edd} \sim 5 \times 10^{34}\,{\rm ergs^{-1}}\eta{\rm M_J} \sim 13\,{\rm L_\odot}\eta{\rm M_J} \tag{6.12}$$

where η takes into account geometry and other parameters, $\eta \sim 1$. Also, a peak temperature can be given:

$$T_{\rm peak} = \left(\frac{L_{\rm peak}}{4\pi R_2^2\sigma}\right)^{1/4} \sim 1.1 \times 10^5\,{\rm K}\eta^{1/4}m_J^{1/4}r_e^{-1/2} \tag{6.13}$$

From that temperature in combination with the Planck law, the thermal spectrum can be calculated.

The following timescales are found:

- Timescale for the prompt flash where the emission is 50% above the peak luminosity:

$$\tau_1 \sim 10\tau_{\rm sh} \sim 2{\rm hr}(r_E/C_s) \tag{6.14}$$

- Bright afterglow:

$$1.4\times 10^3{\rm yr}\,(M_E M_J)R_J^{-3}\left(\frac{T_0}{2,500{\rm K}}\right)^{-4} \leq \tau_2 \leq 10^4\,{\rm yr} \tag{6.15}$$

The peak of this radiation is in the IR.

- By taking into account the convection on a Jupiter sized planet, the timescale to achieve a full-surface isothermal state is

$$\tau_3 \sim 1\,{\rm month} \tag{6.16}$$

6.1.2.1 Observation of Exoplanetary Collisions

In the above cited paper also, observational aspects are discussed. The convection-induced turbulence amplifies magnetic fields, relativistic electrons are accelerated forming shocks (reconnection), and a bright radio flare lasting for months can be observed. Therefore, nearby events (within 3 kpc) should be seen from radio observations. Besides observations in the radio spectrum, simple observations can also be made by photometrically monitoring a huge number of stars over a long period of time. In the B- and V-band,[1] a 2–20% increase of the fluxes is expected for late-type stars. The future GAIA mission could monitor a sample of 10^9 stars for several years and detect several collisional events. IR afterglows have a longer duration, however, the flux contrasts are lower. Therefore, in the IR, the stars should be observed individually using interferometers; the most promising targets would be star-forming regions.[2]

6.1.3 Small Body Collisions

We now address the more frequently occurring collisions of small bodies at early stages of planetary system formation. Can they be observed in exoplanetary systems?

The extension of a planetary system depends on the concentration of protoplanetary particles in the circumstellar disk. Calculations show that a higher concentration leads to a smaller extension: for 0.25 $M_\odot$, the system extended up to 15 AU from

[1] The B band is in the blue contiunuum centered at 442 nm; the V band is centered at 540 nm.

[2] The nearest star forming region is at a distance of 130 pc.

the star; for 0.02 $M_\odot$, the system extended up to 250 AU [235]. In the high mass model, most of the solids are lost in the star.

The star β *Pictoris* is at a distance of 62.9 Ly. In 1983, a circumstellar disk was discovered around this star, by measurements with the Infrared Astronomical Satellite (IRAS). For this star, the following astrophysical parameters are known: spectral type A5 V (maybe also A3 V), age between 20 and 200 million years, mass $1.75\,M_\odot$, luminosity $8.7\,L_\odot$, and radius $1.4\,R_\odot$. The extension of the disk is 1,000 AU.[3] The inner edge of the disk begins at 25 AU. The disk consists of dust grains with sizes below 2–10 μm. Because of the star's radiation pressure, these particles move away from the central star and the age of the disk must be less than 10 million years. Its presence may be explained by impacts of icy *Edgeworth-Kuiper Belt* type objects on large disk bodies. In the inner region (< 10 AU), the dust amounts to 10^{22} g. Numerical simulations to demonstrate this dust production by collisions have been made by Thébault, Augereau, and Beust [243].

Collision at early phases of planet formation are important for the formation of Earth-sized planets. In the well known open star cluster Pleiades, an F5 star HD 23514 was investigated. The presence of small dust particles can be inferred from a strong mid-IR silicate emission feature. However, this emission is very strong and, thus, the existence of so much dust within an AU or so of these stars is not easily accounted for, given the very brief lifetime in orbit of small particles. It can be explained by a collision of massive bodies: the mass of the bodies involved can be estimated as equivalent to the Moon's mass. This mass must be converted, via collisions of massive bodies, to tiny dust particles that find their way to the terrestrial planet zone during the first few hundred million years of the life of many (most?) Sun-like stars (Rhee, Song, and Zuckermann [207]).

Another object studied is BD +20 307. It is a binary system with a very high IR excess. The two objects are similar in size and temperature. The primary shows a weak Li 670.7 nm line. This object is older than 1 Gyr. The large amount of hot circumbinary dust must be from a very large and recent, but very late evolutionarily, collision of planetesimals [262].

Morlok et al. [170] studied dust from collisions in circumstellar disks and the possible similarities to meteorites.

The migration of planets was described already. As it was pointed out by Murray et al. [175], a planet will migrate inward toward the host star when the surface density of planetesimals exceeds a critical value. This would be equivalent to about 0.03 solar masses inside the orbit of Jupiter when compared with our solar system. Also, the migration of planets can lead to catastrophes. The role of giant planets as a shielding against comet and asteroid impacts on smaller planets is important, but there is some controversy too. On the one hand, Jupiter attracts cometary and other small bodies providing some protection to the inner planetary system; on the other hand, it perturbs Kuiper belt objects and these objects penetrate to the inner planetary system. A similar situation is conceivable in other planetary systems.

[3] About 20 times the Sun-Pluto distance.

6.2 Case Studies: Late Type Stars

6.2.1 General Properties

By the term late type star, usually stars later than spectral type G are denoted in astrophysics. Here, we define late type stars as stars of spectral types K and M. On the HRD, they appear at the right hand side of the Sun. As it has been argued already, for studies about habitability, it is reasonable to concentrate on main sequence type stars; we will only briefly mention the giant and supergiant type stars.

In Table 6.1, we give the spectral class, temperature, mass, and approximated main sequence lifetime (MS lifetime) for these and solar-like stars (spectral type G).

Table 6.1 Solar-like and late type Stars

Spectral Type	Mass	Temperature	Approx. MS lifetime
G	1.0	5,800	10×10^9
K	0.6	4,000	32×10^9
M	0.22	2,800	210×10^9

Generally, late type stars are cool stars and their masses are less than the mass of the Sun.

6.2.2 G, K Stars

G stars are solar like and they appear as yellow stars; K stars appear as orange stars. The spectra are dominated by the H and K lines of calcium as well as lines of neutral iron and Ti. Because of the relatively low temperature, also, molecular bands of CN and TiO are visible, especially at the transition from K to M stars. The K type main sequence stars have masses between $0.5 < \mathrm{M}_\odot < 0.8$ and the temperature is between $3,900$ to $5,200\,\mathrm{K}$. Their luminosity ranges from 0.1 to $0.4\,\mathrm{L}_\odot$. Some examples of K stars will be discussed in detail.

6.2.2.1 ε Eri

The star ε Eri is at a distance of 3.22 pc and its mass is 0.75 $\mathrm{M}_\odot$. The luminosity is about 0.3 $\mathrm{L}_\odot$. Its rotation period is 11 d. Due to this rapid rotation, the star generates a strong magnetic field and this results in a high chromospheric activity. Also large starspots have been detected indirectly on its surface. It belongs to a group of variable stars that are known as *BY Dra* stars. It is not likely to find a HZ around this star because it is very young (between 500×10^6 and 1 billion years). Submillimeter observation revealed that the star is surrounded by a dust ring which

is located at the same distance as the Kuiper Belt from the Sun. In the ring, one observes a bright feature which could be a hint of a planet that is accreting matter as well as a dark feature that could be another planet that already has stopped accretion. It is believed that this object is in a stage as has been the solar system 4 billion years ago. So far, the existence of a Jupiter-sized planet and a planet with 0.1 M_J was confirmed. Combining ground based observations with HST observations, Benedict et al. [13] found a perturbation period of 6.85 years, a semi-major axis of $\alpha = 1.88 \pm 0.2\,\text{mas}$[4] and an inclination of 30°.1. The mass of ε Eri b was found by these authors as M=1.55 $M_J \pm 0.24\,M_J$. Attempts to detect this object in the IR were made with the Spitzer telescope, [153] however, with no definite results yet. The star seems to exhibit strong variations. The magnetic activity can be derived from the H and K emission. Gray and Baliunas [80] report that they found a dominant activity longitude on the star as well as temperature variations (15 K) and variation of the granulation (35 m/s). From these variations they predict luminosity variations of 1.2%, which is 10 times the value of the present Sun. Temperature measurements can by made by determining the ratio depths of the vanadium V I 625.183 nm to the Fe I 625.257 lines that has a sensitivity of about 10 K per 1% variation in the depth ratio.

The measurement of the variation of granulation was performed by measuring the bisectors. Bisectors are the lines that half a line profile at equal intensity levels. Since the core of a line originates higher in the atmosphere than the region near the continuum of the line profile, bisectors represent vertical velocity profiles over the range where the line profile has been formed. If there is convection on the surface of a star, line profiles originating from upward moving convective elements will be blueshifted (because the matter approaches the observer) and line profiles that originate in the downflowing areas will be redshifted. For stars, however, we are not able to observe individual up- and downward moving elements but the line profile is a superposition of all the profiles from up- and downward moving elements. This introduces some asymmetry because

- the velocities of the upward moving elements are smaller than those of the downward moving elements,
- the areas of the upward moving elements are larger than those of the downward moving elements.

Therefore, the bisectors are curved ("C"-shaped). For the estimation of the behavior of granulation, 50% to 80% of the mean bisector values were taken for the line 625.3 nm.

If we assume that there exist a planet in the HZ, around that star, then, the predictions are not very promising there. First, the star is too young, so that life had perhaps no chance to have already evolved yet. Even if this would be the case, then we must consider the catastrophic variations on the climate of such a hypothetical planet induced by the host star. As a crude estimate, we can say that the known 0.1% luminosity variations (overall luminosity) over the course of one solar cycle lead to global temperature variations on the Earth's surface of about 0.1 K (that is why the

[4] mas denotes milli arcsec.

solar signal is so difficult to detect). An Earth-like planet in the ε Eri system would be exposed to strong climate variations caused by the shorter and much stronger activity cycle of its host star leading to global climate variations of perhaps more than 1 K every cylce. Moreover, because the star has strong chromospheric activity, these variations in the UV and X-range would extremely influence a planetary atmosphere.

6.2.2.2 τ Ceti, HD 10700

This star is the most Solar-like object among the 30 nearest stars. The spectral type is G8Vp, the luminosity $0.59\,L_{\odot}$, and the mass $0.81\,M_{\odot}$. Its distance is 3.64 pc (11.89 Ly).[5] Around this star, a disk containing dust and cometary like objects has been detected. However, up to now, no planetary objects has been detected. It is not sure whether such objects could have evolved there because the star has a very low metallicity. If there would be any planets, then, they would be subjected to a much more intense bombardment by asteroids and comets than in our solar system. The dust particles (debris) extend out to 55 AU from the star and can be observed in an 850 μm image. It is assumed that the colliding bodies are up to 10 km in diameter and that the mass of the belt is 1.2 M_{Earth}[6] [81].

The age of τ Ceti has been determined 10 ± 0.5 Gy. Why this solar-like star has such a massive Kuiper belt remains to be answered. Obviously, τ Cet has retained a large cometary population.

6.2.2.3 Dust Disks Around Nearest Stars

Within about 5 pc, it is possible to confirm the existence of such dust disks around the central stars. There have been detected five G to mid K-systems namely:

- α Cen
- 61 Cyg
- ε Ind
- ε Eri and
- τ Cet.

The first three of them have companions that can tidally strip disks containing debris and the last two are examples of extreme dust rich systems.

6.2.3 M Type Stars

M stars are cool red stars with a surface temperature less than 3,600 K. Therefore, molecular bands (e.g., TiO) are quite prominent in their spectra. The main sequence

[5] The visual magnitude is 3.49, so it can be easily seen with the naked eye.

[6] The mass of the Kuiper belt is 0.1 M_{Earth}.

M stars (red dwarfs) have a mass of $< 0.5 M_\odot$ and a luminosity $< 0.08 L_\odot$. The M type giant stars (red giants) have higher luminosity ($> 300 L_\odot$) and the M type supergiants have luminosity up to $10^5 L_\odot$.

6.2.4 Proxima Centauri

One example of a red dwarf star is *Proxima Centauri*, which is the nearest star to our solar system at a distance of 4.24 Ly.[7] It is a very faint star of magnitude $11^m.1$, spectral class M5e. Proxima Centauri is at a distance of 13,500 AU from the double star α Cen A, B. It orbits this pair of stars with an orbital period of 500,000 years. Therefore, the system α Cen is in fact a triple system (and Proxima Centauri is also called α Cen C). The period of revolution of the primary components α Cen A, B is 79.92 years. Some physical parameters of the components are given in Table 6.2.

It is clear from these values that the detection of red M-dwarfs at far distances becomes very difficult because of their low luminosity.

According to the Big Bang theory, the universe started from a phase of extreme condensation and extreme temperature. It expanded and during this expansion, it cooled. There was a time span where nuclear fusion lead to the creation of He. Thus, the primordial composition of the universe was H and He with no other elements. Since red dwarfs have a very long lifetime on the main sequence, one would expect to detect some of them with extremely low metal content, but this was not the case so far. It seems that red dwarfs are very common objects in the universe. Considering the 30 nearest stars to our planetary system, 20 of them belong to this category.

Table 6.2 The system α Cen

Component	Mass	Apparent Magnitude	Diameter	Spectral Type
α Cen A	1.1	$0^m.01$	1.227	G2
α Cen B	0.85	$1^m.34$	0.865	K4
α Cen C	0.123	$11^m.05$	0.145	M5e

6.2.4.1 Flare Stars: Catastrophic Bursts

Flare stars are red dwarfs that are variable stars. The variability could be very large, a strong increase in brightness for a few minutes or a few hours can occur. The increase is observed at all wavelengths from the X-rays to radio waves. Flare stares are also called *UV Ceti* variables. This prototype underwent an increase by four magnitudes, which corresponds to a rise in effective temperature about 10,000 K. The flares observed there are characterized by a steep increase and a slower decrease

[7] This distance is 268,000 times the Sun–Earth distance.

Table 6.3 Flare Stars, *V* denotes the visual magnitude

Object	Distance	Sp.class	V	Lum.	Mass	Diam.
EV Lac	16.5 Ly	M3.5	10.30	0.0017	0.16	0.38
AD Leo	16.0	M3	9.4	0.00361	0.2	0.42
UV Ceti, L726-B	8.554	M6e	12.70	0.00005	0.1	0.14
L726-8 A	9.554	M5.5e	12.57	0.00006	0.12	0.15

in the light curve and the greater the increase, shorter the wavelength. Their origin is thought to be similar to the origin of flares on the Sun by magnetic reconnection. The flares are completely unpredictable. Other examples of flare stars are *EV Lac*, *WX UMa*, and *AD Leo*. Some parameters of flare stars are given in Table 6.3. Note that UV Ceti is a binary (period of revolution 200 years).

Flares can be observed by different techniques and in different wavelengths. There are many systems that are binary systems and examples were found where the flaring periodicity is caused by an interaction of a large magnetic structure from both components. Another example shows more of a solar-like behavior, a flaring periodicity that seems to have its origin in the stellar interior. Similar to our Sun, magnetic flux tubes then emerge to the surface and magnetic reconnection is triggered [158].

Another example of a large flare is the observation of a flaring event on the M4 dwarf star *GJ 3685A*. The observations were made simultaneously in different wavelengths with the Galaxy Evolution Explorer (GALEX) satellite Near-ultraviolet (NUV, 175–280 nm) and Far-ultraviolet (FUV, 135–175 nm) and time-tagged photometry, with a time resolution better than 0.1 s showing that the overall brightness in the FUV band increased by a factor of 1,000 in 200 s [210].

6.2.4.2 Planetary Systems in Late Type Stars

Stars of spectral type later than K are cooler than the Sun. Consequently, the HZ must be closer to them than for hotter stars.

On the other hand, cooler stars tend to be unstable and more active as it is seen by the example of the flare, stars. The flares are unpredictable. During the flare, the brightness of the star can double and is of great importance for a HZ there; the emission in the X-rays can be up to 10^4 that of comparable flares on the Sun. This certainly will have deadly consequences for any life in HZ. The exact percentage of how many M-dwarfs are actually flare stars is not known, but it seems that most of them belong to that class.

Buccino et al. [26] estimated the ultraviolet radiation constraints around CHZ. It is well known that UV radiation between 200 and 300 nm is extremely damaging to most of biological systems (at least terrestrial ones). But we also have to take into account that UV radiation was one of and the most important energy sources for biogenesis processes on the primitive Earth.

6.3 Early Type Stars

6.3.1 Properties of Early Type Stars

Early type stars are found on the left hand side of the HRD. Thus, they are more luminous and more massive than late type stars. Because of their larger masses, their evolution on the main sequence of the HRD is much faster and they contribute to the chemical and compositional evolution of galaxies. Because of their larger temperatures, the radiation is mostly in the shorter wavelengths, namely in the blue to UV regions of the spectrum. It is generally accepted that all stars from spectral types O, B, A, and F0–F5 are classified as early type stars. In Table 6.4, the masses, surface temperature, and approximate main sequence lifetimes for these stars are given.

Table 6.4 Some parameters for early type stars

Spectral Type	Mass	Surface Temperature	Main Sequence Lifetime
O	32	35000	10×10^6
B	6	14000	270×10^6
A	2	8100	800×10^6
F	1.25	7000	4.2×10^9

The percentage of early type stars (O–F) is less than 4% of all stars on the main sequence.

6.3.2 Planetary Systems in Early Type Stars

The environment around early type stars is extremely hostile to formation of planetary systems. Observations with the Spitzer Space Telescope[8] have shown that around O stars, planetary formation becomes impossible because of the *photo-evaporation* effect. This effect comes from interaction of radiation pressure with matter. It is well known for comets (where it forces their dust tails to point in the opposite direction of the Sun). If the radiation pressure of a star is strong enough because of the intense radiation of the star, then, the dust and gas particles are expelled out. Therefore, the presence of very hot luminous stars in the so-called star the forming regions certainly has consequences for star formation there, but it is not clear whether it has positive or negative effects. The B stars are also extremely bright and they constitute 0.13% of all main sequence stars. Examples on the sky are *Rigel* and *Spica*. Together with O stars they are detected in OB associations where also molecular clouds are found.

Images taken at 24 μm with the Spitzer Telescope showed a cometary like structure of the dust that is pulled away from the disk about high massive O type stars.

[8] Launched in 2003, it obtains images and spectra in the range 3 μm to 180 μm.

This can also explain the so-called *proplyds* observed in Orion. The evaporation time is given in the range of $1-3\times10^5$ yr (e.g., Balog *et al.*, [7].

Based on the assumption that gas giant planets form rapidly ($\sim 10^2$ yr) due to strong gravitational instabilities in accretion disks, one can estimate what level of EUV radiation can lead to heating and significant atmospheric loss from a planet [113].

Throop and Bally [246] calculated that UV radiation can even stimulate the formation of planetesimals in externally illuminated protoplanetary disks. They assumed an external O or B star that illuminates on a disk. This illumination lead to an increase of the dust-to-gas surface density disk ratio in the disk's interior. Finally, the disk became instable and the formation of planetesimals was accelerated.

Photoevaporation might also be an explanation why migration halting of giant planets occurs. As we already explained, giant planets can migrate toward their central stars because of tidal interactions between the planets and the gas disks of which they were formed. Due to photoevaporation, an excess of systems with planets at or outside the photoevaporation radius can be explained [159].

Class A stars are white or bluish-white and examples are *Vega*, *Sirius*, and *Deneb*. Examples of F stars are *Canopus* and *Procyon*; they represent 3.1% of all main sequence stars.

Habitability seems to be only interesting for F type stars because their main sequence lifetime is several billion years. Also, their radiation environment is less hostile than that of the hotter O, B, and even A stars. Stellar activity could be similar to that of solar-like stars.

Chapter 7
The Solar Neighborhood

In this chapter, we describe the neighborhood of the Sun. Since the Sun rotates around the galactic center, this neighborhood changes and this could have consequences for habitability in the solar system. Perturbations by nearby passing stars could trigger comets from the Oort cloud to penetrate into the inner solar system and the consequences of such impacts have been already discussed. Other effects will be mentioned as well.

7.1 The Sun in the Galaxy

7.1.1 Interstellar Matter

The space between the stars is not empty but filled with the interstellar medium, which contains dust and gas.

The dust component can be observed as a reddening and absorption of starlight. Therefore, the colors of distant stars have to be corrected. The absorption can be detected from stellar statistics and even observations of the galaxy with the naked eye reveal large dark dust clouds. The particles are oriented along the galactic magnetic field and this orientation of the dust particles causes a polarization of starlight. Stars with strong stellar winds at their giant phase in their evolution are the major sources of the dust particles; in cool stars, two emission features are measured at 9.7 micrometers (silicate dust) and 11.5 micrometers (silicon-carbon signature).

The gas component of interstellar matter can be observed either as neutral hydrogen H I at the radio wavelength of 21 cm, or as bright emission regions of ionized hydrogen H II (Hα emission nebulae). The average density of the interstellar medium in the galaxy is about 0.5 atoms cm^{-3}.

The complex molecules that are found in the interstellar matter have been discussed already.

A. Hanslmeier, *Habitability and Cosmic Catastrophes*, Advances in Astrobiology and Biogeophysics, DOI 10.1007/978-3-540-76945-3_7,

7.1.2 The Local Bubble

The local environment of the solar system is different from the conditions found in interstellar matter. The solar system is located in the so-called *Local Bubble*, which is a cavity in the interstellar medium. The Local Bubble was discovered by measurements of starlight reddening [61]. The Local Bubble void was created by star formation processes. A blast wave evacuated a low density region and compressed surrounding molecular clouds; massive OB stars were formed, which is known as *Gould's Belt*. The Sun is moving away from the center of the Gould's Belt and is closest to the Scorpius-Centaurus rim.

The *local interstellar cloud (LIC)* has a density of neutral hydrogen $n(H) \sim 0.25 - 0.30\,\mathrm{cm}^{-3}$.
From pulsar dispersion measures, the uniform component of an interstellar magnetic field B_{IS} near the Sun was found to be $\sim 1.4\,\mu\mathrm{G}$ with correlation lengths $\sim 100\,\mathrm{pc}$. There exists also a specific radiation environment. The sources for these radiation are

- Plasma emission from the Local Bubble interior
- Supernova remnants, they also provide a diffuse soft X-ray background
- Stellar radiation
- White dwarf radiation.

More information about the Local Bubble can be found, e.g., in the review by Frisch [65]. A search of all B stars within a volume of 400 pc diameter that tracks OB associations which hosted the supernovae responsible for the Local Bubble was presented by Fuchs et al. [66]. They found that 14–20 supernovae have gone off since the past 10–15 million years. The solar system has been travelling this region since the last 3 million years.

The supernova explosion of *Geminga* occurred 300,000 years ago and pushed away the interstellar medium, thus, a region that is less dense formed. The supernova remnant is a neutron star seen as a pulsar in the constellation of Gemini at a distance

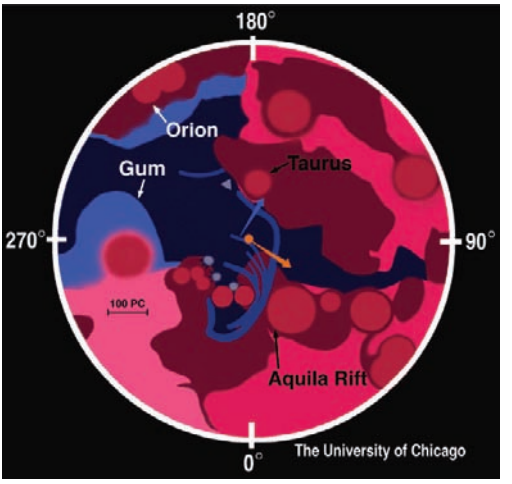

Fig. 7.1 At the center, the present location of the Sun in the Galaxy is shown among different gas clouds. There are nearly spherical bubbles which surround regions of recent star formation. The filaments near the Sun are gas shells. They are remnant from star formation that took place 4 million years ago in the Scorpius-Centaurus Association. Courtesy: P.C. Frisch, C. Wellman, U. Chicago

of 552 Lyr. At present, the position of the solar system is the so-called local fluff or local interstellar cloud where the density is about 0.1 atoms cm^{-3}.

In Fig. 7.1, the present neighborhood of the solar system is shown. Over the past few million years, the Sun has moved through a region of relatively low density gas in the galaxy. Besides that motion, it oscillates with respect to the galaxy's plane with a period of about 66 million years. Clouds of dark gas in the solar neighborhood were investigated by Grenier, Casandjian, and Terrier [82].

7.2 Nearby Stars

7.2.1 Definition of the Solar Neighborhood

The real number of objects near the solar neighborhood (Fig. 7.2) is difficult to determine because there may be many small, not very luminous objects such as brown dwarfs. Kirkpatrick [124] estimates that cool objects like brown dwarfs may outnumber the stars two to one. But because of their small sizes, they contribute only to about 10% of the mass in the solar neighborhood. The sample was restricted to objects within a distance of 25 pc (Fig. 7.3).

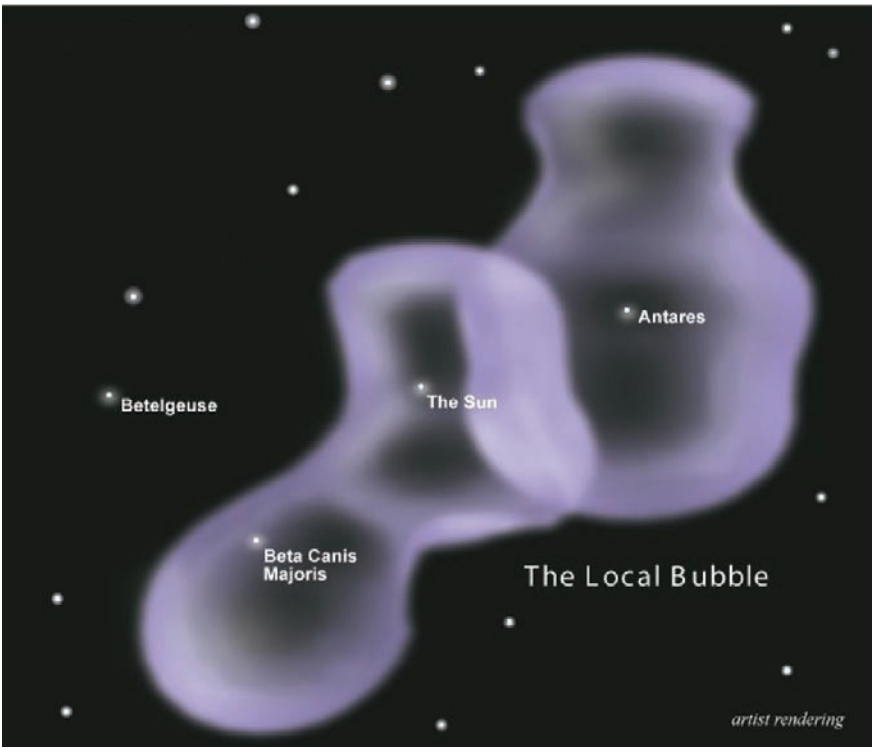

Fig. 7.2 The Local Bubble and the position of the solar system. Courtesy: NASA

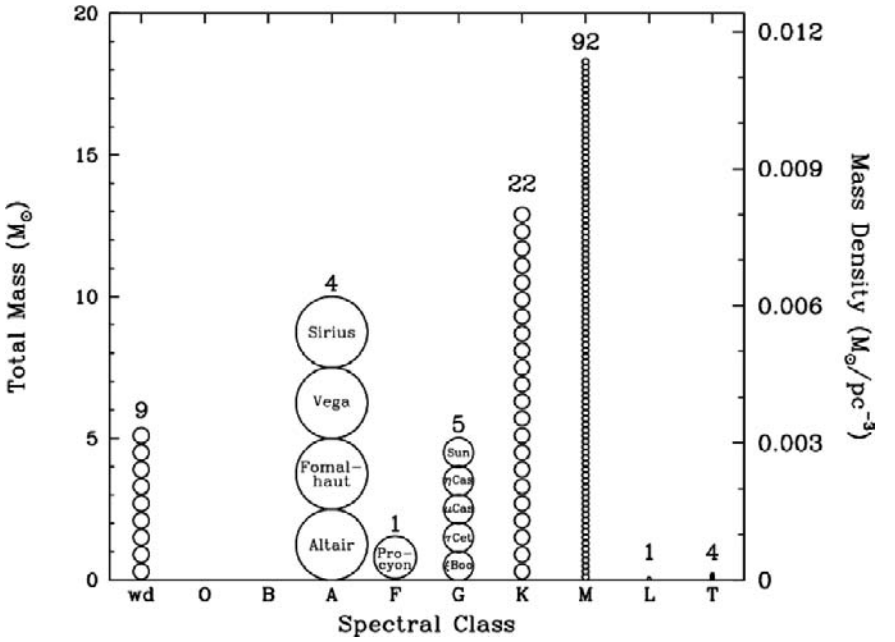

Fig. 7.3 The Solar neighborhood. Adapted from: Kirkpatrick, D., 2001 [124]

7.2.2 Nemesis

In this section, we shortly describe a hypothesis that was first postulated by Whitmire and Jackson [268]. It was found by Raup and Sepkoski [203] that mass extinctions on Earth have occurred with a period of roughly 26 million years. Assuming a solar companion at a distance of 88,000 AU (1.4 Lyr) in a highly eccentric orbit, this object would cause disturbances in the Oort cloud initiating intense comet showers that penetrate into the inner solar system. The companion star would have a mass in the black dwarf range of 0.0002 to 0.07 solar masses.

Though not very luminous in the visible range of the spectrum, this object should be observable in the IR. Dynamical constraints on the mass and perihelion distance of Nemesis and the stability of its orbit are studied in Hills [103]. Muller suggests that the magnitude of that object should be between 7 and 12.[1] From the expected magnitude of such an object and by a comparison with catalogs (optical and IR), Bhalerao and Vahia [16] estimated that the Sun cannot have a companion of mass greater than 44 M_J (0.042 $M_\odot$).

Are there other examples of a "Nemesis" in extrasolar systems? The planet-hosting star HD 3651 has a faint, comoving wide companion HD 3651B (mass 0.033 $M_\odot$, T=790 K (see Burgasser [29]).

[1] http://muller.lbl.gov/pages/lbl-nem.htm

7.2.3 Mass Extinctions by Galactic Clouds

The basic properties of the gas and dust component of our galaxy have already been discussed as well as the great variety of organic compounds that can be found there. In this section, we want to mention another aspect of the interstellar medium.

The solar system rotates about the center of the galaxy and one revolution lasts for about 220×10^6 yr. That is not the only motion of the solar system in the galaxy. The Sun also oscillates vertically through the galactic disk. Because of this motion, it passes from time to time through the galactic plane. This passage through the galactic plane occurs in a cyclic time of about 30–36 million years.

The galactic disk is a region of higher concentration of stars and interstellar medium.

The mass density is proportional to

$$\rho \sim V^2/D^2 \tag{7.1}$$

V is the velocity perpendicular to the galactic plane and D is the thickness of the disk. Data from the HIPPARCOS satellite gives a mass density of

$$0.076 \pm 0.015\,\mathrm{M}_\odot\,\mathrm{pc}^{-3} \tag{7.2}$$

As we have mentioned, the solar system is surrounded by a spherical shell of about 10^{12} cometary or asteroid like objects, the *Oort cloud*. The molecular clouds near the galactic plane can, therefore, interact with this outer parts of the solar system. This was investigated by Rampino [198]. When the Sun passes through the galactic plane, such interactions would be much more frequently. This would induce a quasiperiodicity of large cometary impacts on Earth (see also Stothers [239]). Napier [176] discusses that large terrestrial impacts tend to cluster in discrete episodes, with characteristic separations 25–30 Myr and durations of about 1–2 Myr. As a result of the action of the Galactic tide on the Oort comet cloud through the passage of the solar system in the galactic plane, objects in the Oort cloud are perturbed and penetrate into the inner solar system. This may also happen when the solar system passes interstellar clouds, stars, or encounters the spiral arms.

7.3 Habitable Zones in Galaxies

7.3.1 Supernova Explosions

On the average, a supernova may occur once every 50 years in a galaxy that is like our own system. As we have already exposed, supernovae are important for the cycle of elements in the cosmos (Fig. 7.4), because, during their explosion, elements that are heavier than He are set free to the interstellar medium. These elements

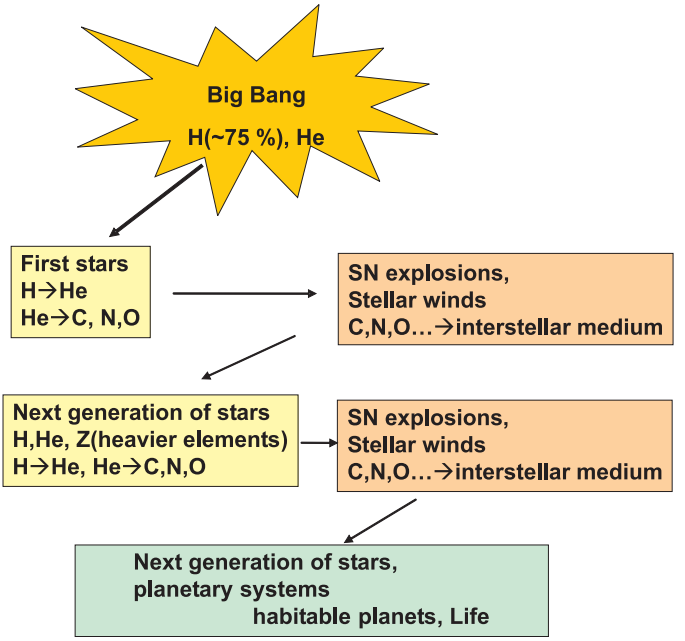

Fig. 7.4 The cosmic cycle of matter

had originated during various fusion processes in the stars before the explosion (the heavier elements were produced fractions of a day before the explosion) and a supernova explosion occurs when a *massive star* with mass $M > 1.4\,M_{\odot}$ collapses under the force of its gravity or a when a *white dwarf* accumulates matter above that limit.[2] The first type of supernovae is called type II (*Type II SN*), the second type I (*Type I SN*).

The energy of $1-2\times10^{44}$ J is set free within a few seconds. A shock wave moves outward and the velocity is in the range of 5–20,000km/s.

There are estimates that a near-Earth supernova, i.e., within a distance of $<$ 100 Lyr would have severe effects on the *biosphere*: the most dangerous are the emitted *gamma rays*. The gamma rays react with molecular N in the upper atmosphere, the nitrogen atom combines with an oxygen atom, and nitrogen oxides are formed. These nitrogen oxides cause a *depletion of the ozone layer*. Thus, there will be less protection of shortwave UV radiation from the Sun and cosmic radiation and the surface will be exposed to a higher radiation dose.

The ozone depletion occurs mainly through the following reaction:

$$NO + O_3 \rightarrow NO_2 + O_2 \tag{7.3}$$

$$NO_2 + O \rightarrow NO + O_2 \tag{7.4}$$

[2] This is also called the Chandrasekhar limit.

Therefore, the net reaction is

$$O_3 + O \rightarrow O_2 + O_2 \tag{7.5}$$

According to Melott et al. [163], an increase of UVB radiation (315–280 nm) in the range of only 10 %–30 % will have lethal effects. Especially, *phytoplankton* will be affected, which is the base of the food chain.

As an approximation, the effect of a nearby supernova explosion can be treated as follows:

- Short-lived gamma-ray blast (300 d) $\sim 2 \times 10^{47}$ erg
- Longer-lived cosmic-ray component (20 yr) $\sim 4 \times 10^{49}$ erg.

Gehrels et al. [69] showed that for a supernova to be strong enough to halve the ozone column density, it must occur within 8 pc. This would double the biologically active UV radiation reaching the surface and produce significant perturbations and mutations. It was estimated that on the average such an event has to be expected at a rate of 1.5 per Gyr.

Ellis and Schramm [55] came to the conclusion that a supernova explosion of the order of 10 pc away could be expected as often as every few hundred million years and could destroy the ozone layer for hundreds of years, letting in potentially lethal solar ultraviolet radiation.

Table 7.1 summarizes the nearby stars that might explode as supernovae in the not too far future.

The most probable candidate is η Carinae (Fig. 7.5), which might explode as a supernova within the next 10,000–20,000 years. This object is classified as a luminous blue variable (LBV). It is a binary system and the masses of the components are $80\,M_\odot$ and $30\,M_\odot$. Several luminosity outbursts of η Carinae were observed. In 1677, *Halley* described it as a faint star of fourth magnitude. Between 1837 and 1856 its magnitude increased and in 1843, it was the second brightest star in the sky (the brightest is Sirius). The distance is between 7,000 and 10,000 Lyr. The star is found within a nebula (NGC 3372).

Table 7.1 Supernova candidates. Scheat and Mira are probably red giants and will therefore, not explode as supernovae

Star	Constellation	Diatnce (lyr)	Type
Scheat	Pegasus	200	Red giant
Mira	Cetus	230	Red giant
Betelgeuse	Orion	1400	Red supergiant
Antares	Scorpius	470	Red supergiant
Ras Algethi	Hercules	550	Red supergiant
Sher 25		25 000	Red supergiant
Eta Carinae		7500	LBV

Fig. 7.5 η Carinae

The peculiar type IIn supernova (SN) *2006gy* in NGC 1260 was the most luminous supernova ever recorded: the peak absolute magnitude[3] was -22^{m} and the total radiated energy was 10^{51} erg. The star had a mass of about $22\,\mathrm{M}_{\odot}$. A similar scenario can also be expected for η Car. The star seems to explode totally without leaving a black hole [229].

What would happen if η Car explodes as a supernova? If we assume a similar energy output as for the superluminous supernova 2006gy, and take into account that the object is about 30,000 times nearer than 2006gy, it can be estimated that (i) ionization effects and ozone depletion will be unlikely to be important (see also Thomas et al. [245]) and (ii) cosmic ray emission should be spread out over 10^4 yr and most probably will also not affect the biosphere.

7.3.1.1 The Endocrine System and Disruptions

Endocrine systems, also referred to as hormone systems, are found in all mammals, birds, fish, and many other types of living organisms. They are made up of

- Glands located throughout the body.
- Hormones that are made by the glands and released into the bloodstream or the fluid surrounding cells.
- Receptors in various organs and tissues that recognize and respond to the hormones.

[3] This is the apparent magnitude of an object at a distance of 10 pc; at this distance, the Sun would be among the fainter stars visible to the naked eye.

Hormones are released by glands and travel throughout the body, acting as chemical messengers.

Endocrine disruptors are exogenous substances (e.g., chemicals) that interfere with the endocrine system and disrupt the physiologic function of hormones. A possible threat, from nearby supernovae is *endocrine disruption* induced by blue light near the peak of the optical spectrum.

How can Supernovae disturb this system? The optical peak of the SN 2006 gy was about 400 nm, in the blue. Such wavelengths can affect the human endocrine system causing mood disorders, melatonin suppression, and even cancer as a consequence of this. Assuming a luminosity of about $3 \times 10^{37}\,\mathrm{Js}^{-1}$ about $0.5\,\mathrm{mWm}^{-2}$ could irradiate the Earth's surface for several months. This is less than the brightness of the full moon, but, as stated, strongly peaked in the blue.

Let us assume an explosion of η Car. For any observer south of -30°, η Carinae is always above the horizon during night. Because of reddening due to interstellar matter, this blue radiation is not a real threat, if a possible supernova occurs for η Carinae. Thus the effect of endocrine disruption could be negligible for η Carinae. However, Brainhard et al. [24] estimated that the biological threshold could be reached for a typical supernova at about 30 pc. This is greater than the above given value for ionization and ozone depletion effects. However, it is also clear that we have to expect such an event roughly once every 20 My on an average.

7.3.2 Gamma-Ray Bursts, GRB

Gamma-ray bursts are bursts of short duration of intense gamma rays lasting from few milliseconds to several minutes. They are the most luminous source in the universe and during the burst, their luminosity even exceed that of a supernova explosion. Usually, there is one GRB/day and they are detected from all directions of the sky.

It is interesting to describe the history of their detection. In the late 1960s, US military satellites watched for Soviet nuclear testing (nuclear test ban treaty). The satellites carried gamma ray detectors on board because during nuclear tests, gamma rays are produced. In 1963, the US Air Force launched the first series of Vela satellites and on 2 July 1967, the Vela 4a,b[4] satellite made the first observation of a gamma-ray burst. However, the first publishing of the results dates to 1973. Brown [25] already pointed out that these bursts might produce ionospheric effects. Because of their high energy, gamma rays are absorbed at relatively low altitude and, thus, they could produce additional ionization there. This ionization might be strong enough to affect VLF radio propagation. A question then arose–where does this radiation come from.

[4] CsI scintillation counters,which are shielded against direct penetration by electrons below 0.75 MeV and protons below 20 MeV.

An event on 14 May 1973 was detected by five separate instruments located in different geomagnetic environments. The direction was far from the Earth and the Sun. Therefore, it was concluded that these bursts must be of cosmic origin [266]. For this event, the total energy was about $\sim 5 \times 10^{-4}\,\mathrm{erg\,cm^{-2}}$ from 11 to 1,500 keV. If it is assumed that the event originated from a nearby star at a distance of 10 pc, this energy corresponds to 5.5×10^{36} erg. This is four orders of magnitude brighter than optical flares from flare stars. It was proposed by Klebesadel et al. [126] that the origin of this event could be extragalactic (Supernova?). This group analyzed 16 short bursts of photons (energy range 0.2–1.5 MeV) between July 1969 and July 1972. The durations ranged from 0.1 s to ~ 30 s and the time integrated flux densities from 10^{-5} to $2 \times 10^{-4}\,\mathrm{erg\,cm^{-2}}$.

The NASA Compton Gamma Ray Observatory with the BATSE instrument[5] was launched in April 1991 and detected more than 2,700 GRB (over nine years) distributed uniformly over the sky and not concentrated toward the galactic plane, which proves their extragalactic origin. In 1999, an afterglow was detected and it was found that the energy is beamed on narrow jets. GRB can be detected only if the jet coincides with our line of sight. The energy output of GRB990123 was estimated to be 10^{43} W. Later on, evidence of Si, S, and Ar were found in the shell of gas surrounding the burst. These elements are typical for supernovae. Thus, GRB are *Hypernova* explosions. A blast wave of material is produced inside a collapsing star travelling outward in an expanding cone. The stellar material collides with gas and dust in the interstellar medium and emissions of X-rays, UV, visible light, and micro waves and radio waves are producing the so-called afterglow. But the Earth must be aligned along the blast axis.

Let us assume that there occurs one supernova explosion/galaxy per 100 years. If there are 100 billion galaxies in the universe, the number of supernova explosions to be expected per year is about 1 billion supernova and 30 supernovae per second in the observable univsere.[6] It is assumed that about one out of 100,000 supernovae produces a hypernova.

7.3.2.1 Gamma - Ray Burst and Mass Extinctions?

It can be estimated that once in the last Gy, the Earth has been irradiated by a gamma-ray burst that occurred in our Galaxy from within 2 kpc. Thomas et al. [244] assumed a typical GRB with a power of 5×10^{44} W at a distance of 2 kpc. If a 10 s $100\,\mathrm{kJ\,m^{-2}}$ burst is assumed that penetrates to the stratosphere, then a globally averaged ozone depletion of 35% reaching a maximum of 55% would be the result. This significant disturbance could be persistent for at least 5 years. A 50% decrease in ozone column density leads to three times the normal UVB (280–315 nm) radiation dose on the surface. At this wavelength band, living organisms on Earth are very sensitive to UV radiation.

[5] Burst and Transient Source Experiment.

[6] For simplicity the age of the universe is assumed to be 10 billion years.

The Ordovician period started at a major mass extinction event, the *Cambrian-Ordovician extinction* event at about $488.3 \pm 1.7 \times 10^6\,\mathrm{y_r BP}$ and at its end another major extinction event occurred about $444.7 \pm 1.5 \times 10^6\,\mathrm{y_r BP}$. Sixty percent of marine genera were wiped out. One hypothesis to explain mass extinction in the late Ordovician is a GRB event. This could have led to a depletion of the ozone layer and intense solar UV radiation would have reached the Earth's surface. Also, a global cooling could have been triggered leading to a glaciation. One or two GRB may happen in our galaxy every $10^9\,\mathrm{yr}$ [163]. It can be estimated that a 10-second GRB would be sufficient to cause such a mass extinction.

7.4 Catastrophes and Habitability in Galaxies

7.4.1 Galactic Collisions, Starburst Galaxies

Galaxies are found in galaxy clusters. Our galaxy belong to a group of galaxies which is called the Local Group. The average distance between galaxies in a galaxy cluster is only about ten times their diameter.

Let us consider our galaxy and its neighboring galaxy, the Andromeda galaxy. The diameter of our galaxy is about 10^5 Lyr; the distance of the Andromeda galaxy is about 20 times this value. Similar proportions between galaxy diameters and their mutual distances are found for other galaxies and, therefore, collisions between them are not unlikely and are observed quite often. The role of galaxy collisions in the early universe and at present and its influence on galaxy formation and evolution and on star formation was reviewed by Struck [240].

Starburst galaxies are a result of galactic collision. There are many examples of irregularly shaped galaxies. These galaxies result from disruption by the gravitational pull of neighboring galaxies. In normal galaxies, the *star formation rate* is about $3\,\mathrm{M_\odot/yr}$. In starburst galaxies, this value can be more than $100\,\mathrm{M_\odot/yr}$. One of the most famous examples of a starburst galaxy is M82, which can even be seen with amateur telescopes. This galaxy experienced a close encounter with the larger M81. In these galaxies, we find concentrations of dust and gas and many newly formed massive stars that ionize the surrounding clouds and create so-called *H II regions*. Since there are many massive stars that evolve rapidly, supernova explosions occur frequently and the expanding remnants also interact with the interstellar gas. This triggers star formation again. The starburst phase comes to an end as soon as the available gas is consumed. Starburst activity usually lasts for only about 10^7 years.

As it was already discussed in the chapter about our galaxy, about 90% of the mass of most galaxies is dark matter that can only be detected by its gravity. In many galaxies, there is evidence of the existence of a SMBH in their center. These SMBHs may cause high activity in the nuclei of galaxies and they are known as AGNs (active galactic nuclei) SMBH.

7.4.1.1 Collisions in the Andromeda Galaxy?

Is there a chance that our galaxy might collide with another galaxy?

Let us recall that our Galaxy is a member of the Local Group of galaxies to which also the Andromeda Galaxy, M31,[7] and others belong to. Thus, we should try to answer the question whether collisions in the Local Group are likely or whether there are indications that such collisions happened or will happen.

Detailed studies of M31 revealed an outer ring of star formation of about 10 kpc diameter. Its center does not coincide with the nucleus of M31. Optical and radio observations also showed that the outer disk of the M31 galaxy is warped and in its halo, numerous loops and ripples have been found. Recent observations revealed a second inner dust ring with about 1.5 kpc diameter and also offset center. These rings can be interpreted as *density waves* propagating in the disk. Numerical simulations can explain the presence of the two rings. They result from a head-on collision of M31 with another galaxy, most probably M32 (see Fig. 7.6). This event took place 210 million years ago [22].

M32 belongs to the Andromeda system and its diameter is 8,000 Ly. It contains about $10^8\,\mathrm{M}_\odot$ in its central region, which is comparable with the density found near the center of the Andromeda Galaxy: the central density is $5,000\,\mathrm{M}_\odot\,\mathrm{kpc}^{-1}$. Its total mass is about $3\times10^9\,\mathrm{M}_\odot$.

Numerical simulation were also applied to study the evolution of gas and stellar matter caused by the collision of the galaxy pair *NGC 3395/3396* [134].

Fig. 7.6 The Andromeda Galaxy with M32 (North of M31, just at the outskirts of the spiral arms of M31. Courtesy: J. Lanoue

[7] The letter M denotes the Messier Catalogue of objects (star clusters, galaxies, nebulae).

7.4.1.2 Collision Between our Galaxy and the Andromeda Galaxy

Numerical simulations give an even more impressive scenario. Our galaxy is likely to collide with M31. As the galaxies approach each other, they keep their spiral shapes up to the point of impact, where so-called “tidal forces” of gravity result in the formation of long plumes of stars, gas, and dust, called *tidal tails*. The centers of each galaxy then merge into one remnant core. Cox and Loeb [46] used an N-body-hydrodynamic simulation to determine what would happen, especially to the solar system and the Sun. They find that the collision will occur before the end of the Sun’s main sequence lifetime, i.e., within the next 4.5 billion years. There might be two close encounters. During the first encounter, a 12% probability is given that the Sun will leave its current position (Fig. 7.7) and after the second encounter there will be a 30% probability that the Sun will be pulled to the extended tidal material. After the encounters, the most likely position of the Sun will be at a distance of 30 kpc from the galactic center (this is three times its current distance). There is even a small probability that the Sun will be no longer part of the Galaxy but will belong to the Andromeda Galaxy. Dubinski [53] gave a nice review about that catastrophic event. Recent observations show that M31 has about twice the mass as our galaxy.

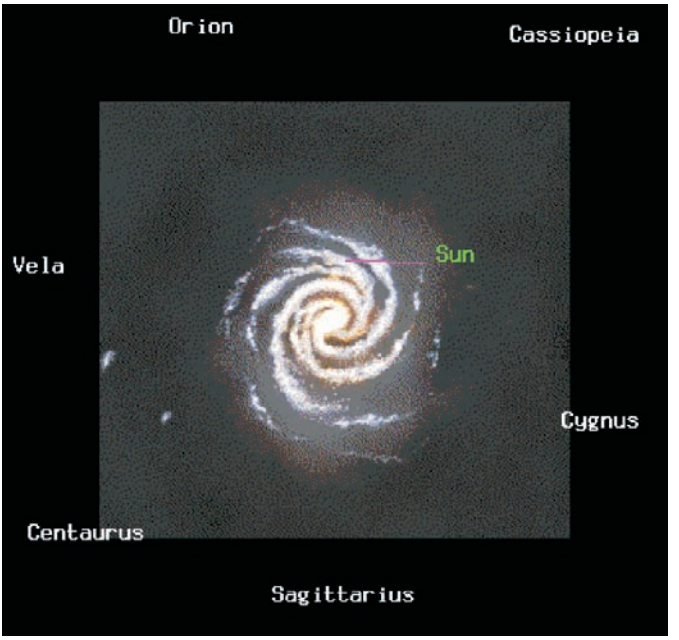

Fig. 7.7 The location of the Sun and the solar system in the Galaxy. After the collision with the Andromeda galaxy it will be at least three times more distant to the galactic center. Image: B. Smith

7.4.1.3 Habitability and Galactic Collision

The above described collision between our galaxy and the neighboring Andromeda galaxy will occur most likely in 3 billion years. Observers on Earth would then probably see two "milky ways" on their night sky. Again, we stress that during galaxy collisions the stars do not collide because of the large distances between them, however, the clouds collide and interact. This changes conditions for habitability dramatically.

- Star formation also in the vicinity of the Sun will be greatly enhanced, thus, rapidly evolving massive stars form and supernovae will explode in our surroundings, which would have deadly consequences for life on Earth.
- Due to the enhanced density of stars and interstellar matter (from both merging galaxies), perturbations in the Oort cloud are more probable and cometary showers that penetrate into the inner solar system will be induced. The impact rates, therefore, strongly increase.
- The conditions of habitability on Earth at that time would be hostile, even without the galactic collision, because the Sun continuously increases its luminosity and surface temperatures on Earth will be above the boiling point of water.
- We may speculate that the zone of habitability would have spread out to Mars and due to cometary impacts induced as a consequence of the galactic collision, his water reservoir might become much larger and maybe habitable conditions could develop there.

Similar considerations can be made for other extrasolar planetary systems. Galactic collision may lead to the emergence of new HZ and destroy existing ones.

Galaxy collisions occurred at a frequency of 10–100 times higher in the early universe because matter was distributed over a smaller space.

7.4.2 Active Galactic Nuclei and Habitability

The nucleus of a galaxy is the central region where the concentration of stars is high and mostly spherical. The nucleus of our galaxy cannot be seen in visible light because of dust clouds. We define the nucleus of a galaxy as active when it emits more energy than is expected from the radiation of stars.

The explanation for the activity in the nucleus is as follows:

- A SMBH (Supermassive Black Hole) in the center
- An accretion disk around this object
- From the accretion disk, matter falls toward the black hole and gravitational energy is released.

Gravitational energy is released when matter falls toward the black hole from the disk and we can estimate that energy:

Assuming that $m = 1 M_{\odot}$ falls into a SMBH of $M = 10^9 \, \mathrm{M}_{\odot}$ from infinity to $R = 3 \times 10^9$ km, then, the gravitational energy

$$E_{\text{pot}} = -\frac{GMm}{R} \tag{7.6}$$

is liberated. We obtain $E = 10^{47}$ J. This is of course a very rough approximation because this energy is not 100% converted into radiation, but is dissipated also due to friction in the accretion disk.

About 1/10 of AGNs show a diametrically opposed pair of energetic jets in which particles are ejected from the core at velocities close to the speed of light.

7.4.2.1 M 87

A well-known example for such a jet is the galaxy *M87*. This giant elliptical galaxy is the dominant object in the Virgo cluster of galaxies. Its mass is estimated to about $10^{12}\,M_\odot$ and its distance is about 60 million light years. The jet can be only observed when the center of the object is not overexposed. It was first observed in 1918. The extension of the jet is about 8,000 lyr. The light is strongly polarized, which is typical for synchrotron radiation. The image given by Fig. 7.8 was taken with the Hubble Space Telescope and shows many knots and other details. In 1966, *Arp* detected another jet in the opposite direction, however, this jet is less conspicuous.

M87 can be also classified as radio galaxy and even gamma rays were detected. The intensity of these rays can vary over few days. This means that the source of this radiation must be compact and the only possible source is a SMBH there. The mass of this is estimated to about $3 \times 10^9\,M_\odot$.

Fig. 7.8 The galaxy core of M87 with its jet. Credit: Hubble Space Telescope

A comparison of the jet in the radio and in the optical region of the electromagnetic spectrum showed that on large scale, there is a great similarity, however, on scales about $0.1''$, the optical emission is more concentrated in knots, the inner jet knots show superluminous velocities,[8] which vary on time scales of 1 year.

During 2002, HST and Chandra monitoring of the M87 jet detected a dramatic flare in knot HST-1 located $\sim 1''$ from the nucleus. As of late 2004, its brightness has increased fifty-fold in the optical band (see, e.g., Perlan et al. [188] or Harris et al. [93]).

7.4.2.2 Blazars, Seyfert Galaxies and Quasars

A radio galaxy emits radio frequencies from relativistic jets. A unified model of these types of active galaxies explains their differences based on the viewing angle of the observer. The orientation of the energetic jet to the line of sight is crucial for the appearance of active galaxies.

If such a relativistic jet directly points in the direction of the Earth, the object is called a *blazar*.

Active galaxies that emit high-energy radiation in the form of X-rays are classified as *Seyfert* galaxies or quasars, depending on the luminosity. *Quasars* are quasistellar objects that show high redshift. Due to the expansion of the universe, there is a linear relation between the redshift z or the radial velocity v, and the distanc d of an object.

$$v = dH \qquad v = cz \tag{7.7}$$

The larger the redshift (radial velocity) the more distant the object. H is the *Hubble constant*.

Because of their high redshift values, Quasars must be at enormous distances and cannot be explained by stars.

7.4.2.3 Our Galaxy—Always Inactive?

The fraction of galaxies observed in the nearby universe with active galactic nuclei is related to the duty cycles of their supermassive black holes. It is estimated that about 40% of massive galaxies have active nuclei. Therefore, the typical lifetime for an active nucleus is in the order of 2×10^8 yr.

This would mean for our galaxy that its nucleus was active about 40 % of its past 10^9 years. In its active state, the luminosity of its nucleus would be about $10^{44}\,\mathrm{erg\,s^{-1}}$ in the 2–10 keV X-rays [86]. This energy is equivalent to the energy of a typical supernova in less than a year. The energy expected on Earth would be (for the X-ray flux) $130\,\mathrm{erg\,m^{-2}\,s^{-1}}$ assuming no absorption. This is about 20 times the typical flux from the Sun and corresponds to the peak flux of an M-class X-ray solar flare. It can be shown that at lower energies, there is a rapidly increasing

[8] These superluminous velocities can be explained by a projection effect.

absorption by interstellar clouds. Taking absorption into account, the flux during an AGN phase of our galaxy should be comparable to the value for the present Sun (X-ray emission).

Could our galaxy become a more active galaxy? There are objects which have much more violent core activities: BL Lacertae objects, e.g., have gamma ray luminosity up to $10^{49} - 10^{50}\,\mathrm{erg\,s^{-1}}$. However, at present, no one can answer the question whether such a violent phase has occurred in the past or will occur in the future of the Galaxy's evolution.

Another important aspect of AGN is their particle emission. Clarke [40] made a very simplified estimation and concluded that the particle radiation fluxes at Earth would increase by a factor of 100 during an AGN phase. This would certainly have an influence on the ozone layer and cause increased UV radiation at the surface. The weak point is to take into account the deflection of the charged particles by the Galaxy's magnetic field.

7.4.2.4 Habitability and Active Galactic Nuclei Phase

Summarizing the influence of an AGN phase on the habitability, we can state the following:

- Radiation: possibly in the order of solar radiation, therefore, no influence.
- Particles: possibly more important, but uncertain.
- Maybe more intense phases in early states of the Galaxy's evolution.

When two galaxies merge (see chapter 7.4.1), a starburst at their common center is initialized as well as the two black holes will merge. This leads to emission of gravitational waves, which transport angular momentum. Black hole mergers can increase the spin of a black hole and also lead to a spin flip, which was observed in radio galaxies. Simulations [147] show that as a consequence of a gas-rich galaxy merger, a black hole with an accretion disk can be ejected and appear as an off-center quasar being observable for several 10^7 yr. Mayer et al. [160] made hydrodynamical simulations of the merging of two massive galaxies with SMBHs. They showed that a SMBH binary will form and shrink and finally a burst of gravitational waves will be emitted. They found these processes to occur on a timescale of 1 million years.

Chapter 8
The Search for Extraterrestrial Life

The search for extraterrestrial life is one of the most fascinating human endeavors. Many institutes and institutions worldwide are involved in this. On the Web, many sites can be found that contain educational information as well as recent research results: the NASA astrobiology institute @SETI (search for extraterrestrial intelligence) is one such example. The search for extraterrestrial life and habitability is closely related topics and we will also synthesize our discussed topics about possible catastrophes with these.

8.1 Life Based on Elements other than Carbon

8.1.1 Silicon-Based Life

So far, we have discussed life that is based on carbon compounds. Carbon has special chemical properties that are important for life and they were given in the first chapter of this book. The question is whether there are other chemical elements with similar properties.

In 1891, J. *Scheiner* speculated on life based on silicon. Two years later, this idea was proposed by E. *Reynolds* [205]. Due to the heat stability of silicon, life based on silicon compounds would exist at high temperatures. About 30 years later, *Haldane* even suggested that silicon-based life may be found deep inside a planet, where molten silicates exist.

Let us briefly discuss silicon biochemistry. Silicon (latin silicium) has the symbol Si and atomic number 14. It is the eighth most common element in the universe and it can be found in dust, planetoids, and planets, being the second most abundant element in the Earth's crust ($\sim 26\%$ of the crust's mass). The density is $2.33\,\mathrm{g}\cdot\mathrm{cm}^{-3}$, the melting point 1,687 K, and the boiling point is 3,585 K. Silicon has the same structure of the outer electron orbitals as carbon. Similar as it is for carbon, the four outer electrons can be donated and shared with other elements. The silicon *silane* SiH_4, for example, is the analogue to the compound methane, CH_4 . Atoms of oxygen and silicon can form polymers that are called *silicones*.

A. Hanslmeier, *Habitability and Cosmic Catastrophes*, Advances in Astrobiology and Biogeophysics, DOI 10.1007/978-3-540-76945-3_8,

Silicon is able to maintain complex molecular structure similar to Carbon. As it has been outlined in the first chapter of this book, long chain molecules form the basic structures for life and from these points of view, life based on silicon compounds is conceivable.

However, let us discuss the different aspects of life: one important aspect is providing the cells with energy via respiration. When carbon is oxidized, it becomes a gas, CO_2, that is easily removed from the body. The oxide of silicon, SiO_2, is not so easily removable. Silicon dioxide or silica is found in nature as sand or quartz and is the principal component of most types of glass. Silicon monoxide was detected in dense interstellar clouds in 1971 at frequencies around 42.8 GHz.

Another property of life is the collection and storage of energy. In carbon-based life, energy storage occurs via carbohydrates. These are oxidized by enzymes.

Silicon chemistry in interstellar clouds and the formation of SiO was discussed by Langer and Glassgold [135].

There is no evidence of silanes or silicones from astronomical observations. Such substances should be the base for silicon-based biota. Another argument against silicon-based life is that silicon is less abundant in the universe than carbon. However, silicon may have played an important role in the chirality. Without chirality, the ability of biomolecules to recognize specific substrates would be crippled, ultimately limiting the number of different reactions available and achievable by a silicon-based system.

8.1.2 Life Based on Ammonia

It was suggested in 1954 by *J. Haldane* [87] that a biochemistry could be conceived in which water is replaced by the solvent of liquid ammonia. There are a number of analogues between water and ammonia. For example, methanol (CH_3OH) has methylamine (CH_3NH_2) as analogue. This is of course of great interest in order to speculate about a biological evolution on ammonia-rich planets such as the gas giants and some of their satellites.

Liquid ammonia can dissolve most organics and also dissolve elements such as sodium, magnesium, and aluminium, so, there is a further similarity to water. At a pressure of 1 atm, ammonia is liquid at -44°C; raising the pressure to, e.g., 60 atm ammonia boils at 98°C. Such pressures are easily achieved on Jupiter or Venus.

Problems are that the hydrogen bonds that exist between the ammonia molecule are much weaker than those in water. Ammonia would be less able to concentrate non polar molecules through a hydrophobic effect. This is essential because prebiotic molecules should be held together to allow the formation of a self-reproducing system.

8.1.3 The Role of Solvents

In our short discussion about biochemistry, we have outlined the importance of solvents in life. To be useful, any solvent must remain liquid within a large range of temperatures. Otherwise, variations in conditions on a planet or satellite of a planet

will freeze or boil the solvent and living organisms will be destroyed. Besides its ability to dissolve other chemical compounds, another important property of a solvent is its ability to help to regulate the temperature of organisms. Solvents also help to transport nutrients and to carry off wastes.

Let us discuss three solvents under pressure conditions found at the surface of Earth:

- Water: the temperature range for water to remain liquid is 100°C.
- Ammonia: the temperature range for liquid ammonia is from −78 to −33°C.
- Methyl alcohol: the temperature range for liquid methyl alcohol is from –94 to +65°C.

For regulating the temperature, the heat capacity is the relevant parameter. Water has a high heat of vaporization; a living cell can respond to a temperature increase by vaporizing just a small amount of water. The heat capacity of water is 1 calorie[1] per gram-degree, the heat capacity of ammonia is 1.23 calorie per gram-degree and that of methyl alcohol is 0.6 calorie per gram-degree. To vaporize, water requires 595 calories per gram, while ammonia needs 300 calories, and methyl alcohol only 290. Thus, water seems to be ideal for temperature regulation. Mammals have a complex brain and precise temperature regulation allows to function this complex brain and cell systems properly.

To form aggregates of organic compounds, the surface tension of a solvent is relevant and for water, this parameter is twice that of ammonia and three times that of alcohol.

Thus, water is much more appropriate for life regarding its heat capacity, heat of vaporization, surface tension, and dissolving capability than other substances. Another important property of water is that it extends as it freezes. The cells of organisms that freeze will rupture. Putting the crew of an interstellar spaceship on ice would not help since humans consist mainly of water and the cells are destroyed by freezing. Ammonia does not expand upon freezing. Astronauts whose life chemistry depend on ammonia as solvent could easily be frozen and then waked up at arrival on some distant stellar system.

Another aspect has to be considered: as we have mentioned, stars produce UV radiation. Water molecules will be dissociated under intense UV radiation and the free oxygen contributes to ozone production. A planet with oceans consisting of ammonia does not produce such a shield.

Life based on ammonia or methyl alcohol will be in some respects more robust and better protected against cosmic catastrophes than life based on water. However, it seems that complex life cannot be based on these because an effective regulation of temperature of the organisms is needed.

8.1.4 Boron-Based Life

Like carbon and silicon, boron has a strong tendency to form covalent molecular compounds. Boron forms a lot of different structural varieties of hydride; in

[1] 1 calorie = 4.1868 J.

Table 8.1 Cosmic abundance of elements. ppm denotes parts per million in mass

Element	ppm	Element	ppm
Hydrogen	739,000	Nitrogen	950
Helium	240,000	Silicon	650
Oxygen	10,700	Magnesium	580
Carbon	4,600	Sulfur	440
Neon	1,340	Boron	10
Iron	1,090	All others	650

these the boron atoms are linked indirectly through hydrogen bridges. Boron forms bonds with nitrogen that are somewhat like the carbon–carbon bond–two electrons from the nitrogen being donated in addition to the covalent electron sharing. Boron-nitrogen compounds largely match the chemical and physical properties of alkanes.

Since Boron is extremely rare in the universe, life based on Boron would not be very probable (Table 8.1).

8.1.5 Nitrogen- and Phosphorus-Based Life

Nitrogen can form long chains of molecules at low temperatures with a liquid solvent such as ammonia (NH_3), or hydrogen-cyanide (HCN).
Phosphorus can also form long chains with C, N, and Si.

8.1.6 Sulfur-Based Life

Also, sulfur is able to from long chains of molecules. On Earth, there are several examples of life that use sulfur as an alternative to oxygen. Sulfur-reducing bacteria produce their energy that is needed for their metabolism by reducing sulfur to hydrogen sulfide. Sulfate reducing bacteria[2] use sulfate (SO_4) as an oxidizing agent, reducing it to sulfide.

A well-known bacterium is Salmonella; they produce hydrogen sulfide. Hydrogen sulfide is structurally similar to water and it dissolves in water:

$$H_2S \rightarrow HS^- + H^+ \tag{8.1}$$

Iron and sulfur based biospheres may play a role on planets like Mars. From different Mars missions, it is established that this planet must have had an early warm and wet climate. If the greenhouse gas CO_2 was at higher concentration at that time, widespread carbonate rocks such as limestone should be found. It has been proposed that on Mars there was a sulfur cycle that was analogous to the carbon cycle on Earth and that SO_2 was the main greenhouse gas. Moreover, sulfur dioxide is a much more powerful greenhouse gas than carbon dioxide [88].

[2] Rotten egg odour indicates the presence of sulfate-reducing bacteria in nature.

The dense clouds of Venus may also provide some refuge for microbial life (whether this life originated on Venus or was brought in by meteorites from Earth). Sulfur allotropes like S_8 might be used as an energy converting pigment by such lifeforms there [219].

8.2 The Gaia Hypothesis

8.2.1 Biota and Their Influences on the Environment

Biological processes are crucial factors in the Earth's surface geochemistry. Let us give a few examples:

- The role of green plants in limiting atmospheric CO_2.
- There are several substances in the Earth's atmosphere that are biogenic or biologically controlled. Some examples are CH_4, N_2, NH_3, NH_4^+, N_2O, NO_2, H_2, SO_4, ...
- Photosynthesis, aerobic and anaerobic metabolism, denitrification, nitrogen fixation, bacterial sulfate reduction, iron oxidation, mineral dissolution by CO_2-rich groundwater, silicates, phosphates...

Thus, there is a strong connection between

biota → biologically induced changes → physical environment.

An example: changes in the albedo due to growth or decay of vegetation. A growth of vegetation decreases the albedo, the surface heats stronger.

8.2.2 Gaia

In 1974, J. Lovelock formulated his Gaia hypothesis (see, e.g., Lovelock and Margulis [149]). The main statement is that "the climate and the chemical composition of the Earth's surface are kept in homeostasis[3] at an optimum by and for the biosphere." The definition of homeostasis is as follows: the property of either an open system or a closed system, especially a living organism, which regulates its internal environment so as to maintain a stable, constant condition.

Basically, the idea is that there is an interaction between nonliving and living systems on Earth. All living things have a regulatory effect on the Earth's environment. Because of these interactions between living and nonliving systems, the Earth is more than just a sphere of rock with a thin layer of atmosphere and ocean. The Earth is regulated by *Gaia*, which means an "entity" that guarantees the optimal conditions for life on it. All living things on Earth from bacteria to humans to whales

[3] The Greek homos means equal, histemi means to stand equally.

are part of Gaia. This implies also that all these beings are important for Gaia. Biological mechanisms control the temperature, pH, and oxidation states in the Earth's atmosphere. Life forms on a planet act with their environment becoming a self-regulating system. This system not only includes organisms but also near-surface rocks, the soil, and the atmosphere.

The biological modulation of the Earth's atmosphere was studied in the paper of Margulis and Lovelock [154].

The Gaia hypothesis was reviewed by Kirchner [123].

8.2.3 How to Test Gaia?

The *Daisyworld* model is a very nice mathematical model and was defined by Watson and Lovelock [260]. This simplified model describes a planet on which temperature is controlled only by the albedo. The albedo is determined by the color of daisies growing on its surface. Black and white daisies are assumed. Both should have the same growth response to temperature. Black daisies are assumed 10° C warmer because they absorb more of the incoming radiation. Therefore, they lower the albedo and warm the surface; the opposite is true for the white daisies. Consequently, the surface temperature is regulated by the growth of the daisies. More black daisies lead to higher surface temperatures.

Let us consider the ice ages. During glacial periods, biological processes produce less CO_2 and the planet, thus, became even colder. This is an example of a positive feedback mechanism.

Also, the climate change could be seen in context with the Gaia concept. This is also reflected in the Amsterdam declaration, 2001:

Four climate change programs were made:

- International Geosphere-Biosphere Programme (IGBP)
- The International Human Dimensions Programme on Global Environmental Change (IHDP)
- The World Climate Research Programme (WCRP)
- The international biodiversity programme DIVERSITAS.

The main conclusions in the Amsterdam declaration are[4]:

- The Earth-System behaves as a single, self-regulating system comprised of physical, chemical, biological, and human components. The interactions and feedbacks between the component parts are complex and exhibit multi-scale temporal and spatial variability.
- Earth-System dynamics are characterized by critical thresholds and abrupt changes.
- In terms of some key environmental parameters, the Earth-System has moved well outside the range of the natural variability exhibited over the last half a million years at least.

[4] http://www.sciconf.igbp.kva.se/Amsterdam_ Declaration.html

8.2.3.1 Gaia, Habitability, and Catastrophes

According to the Gaia hypothesis, the terrestrial surface, biosphere, oceans, and atmosphere is seen as a system of great complexity in which the components strongly interact.

Life and habitability on other worlds, therefore, should be seen in the context of such a concept and not only isolated factors. We have discussed several scenarios that would seriously affect the habitability of our planet. The question is how can the complex system described above react and how can habitability be preserved or even vanish.

If life exists on Mars, the atmosphere of Mars should have developed similar anomalies in temperature, composition, or pH and if no life would be detected on that planet, this could be regarded as a confirmation of the Gaia hypothesis for Earth [155].

8.3 The Future of Extrasolar Planet Finding

8.3.1 Planned Satellite Missions

8.3.1.1 GAIA

The ESA mission GAIA (Global Astrometric Interferometer for Astrophysics) is planned to be launched in December 2011 and the end of the mission is scheduled for 2020. The satellite will be positioned around the Lagrangian point L2, which is located 1.5 million km from the Earth in the anti-Sun direction. At such a position, there are no Earth eclipses. However, because L2 is not a stable point, the spacecraft will orbit about it with a period of about 180 days.

GAIA consists of three modules. The instrument module contains of a dual telescope based on a three-mirror design. Beam combination occurs in a common focal plane. The common focal plane is large with an array of 106 CCD. With that instrument, it is possible to do astrometry (accurate measurements), photometry (continuous spectra in the band 320–1000 nm), and spectrometry (high resolution grating, narrow band 847–874 nm).

The scientific aims of this mission will include a three-dimensional map of our Galaxy and to that purpose, high, accurate positions and radial velocity measurements will be made. The detection and orbital classification of tens of thousands of extrasolar planetary systems is expected. Within 200 pc (650 light years) of the Sun, it is expected to detect 300,000 objects of solar type main sequence stars. The stars are brighter than magnitude 13 and with spectral types earlier than K5.

The methods to detect extrasolar planets with GAIA will be (a) astrometric methods and (b) transit methods. The astrometry method is based on the movement (wobble) of the host star caused by the planets surrounding it. This requires precise measurements over long time spans. By this method, long period planets will

be discovered. By the photometry method planetary transits of Jupiter-size planets around Sun-like stars down to magnitude 14 will be detected. GAIA will also observe known planets around bright stars. This will increase the knowledge about these planetary systems. The science goals of the GAIA mission for exoplanets are as follows:

- Confirm the existence of Jupiter sized planets that have been detected by radial velocity measurements by astrometric methods.
- Observe M dwarfs for such planets.
- Measure the inclinations of the orbital planes.
- Estimate the planet masses.
- Observations of nearby stars which will be further investigated by the future ESA/NASA Darwin/TPF mission.

In total, GAIA will measure the positions of about 10^9 stars in the Galaxy and other members of the Local Group with an accuracy down to 20 μas.[5]

8.3.1.2 KEPLER

The NASA project KEPLER will also search for Earth-like planets. A large sample of stars will be surveyed. The KEPLER telescope, consists of a 0.95 m Schmidt telescope, which is mainly used as a photometer. It has a very large field of view—105 square degrees. This corresponds to a diameter of 12 degrees. The photometer itself consists of an array of 42 CCDs, each 50 × 25 mm with 2,200 × 1,024 pixels. The CCDs are readout every 3 s to prevent saturation. Only the information from the stars brighter than $m_V = 14$ are recorded. Note that the CCDs will not take pictures in the classical sense but the images will be defocussed to 10 arc seconds. By that method, the photometric precision is improved. The data are integrated for 15 minutes.

Thus, a number of stars can be observed simultaneously. KEPLER will observe the same star field for the entire mission and monitor continuously and simultaneously the brightnesses of more than 100,000 stars seen in that field (Fig. 8.1). The planned lifetime of the mission is 4 years. With high precision photometry, it will be possible to detect planetary transits and measure even transits of Earth-like planets. A central transit of the Earth crossing the Sun lasts for 13 h when observed from outside the solar system. Thus, observing brightness variations on this timescale for stars older than 1 Gyr would be a strong hint that a planetary transit has occurred there. Seventy-five percent of the stars older than 1 Gyr are less variable than the Sun on the time scale of transit.

From statistics, it can be expected when taking into account only orbits with four transits in 4 years and assuming that for solar-like stars, planets are common (R_e denotes the Earth's radius):

- About 50 planets if most are the same size as Earth ($R \sim 1.0\,R_e$) and none larger,
- About 185 planets if most have a size of $R \sim 1.3\,R_e$,

[5] $1\,\mu as = 10^{-6}$ arcsec.

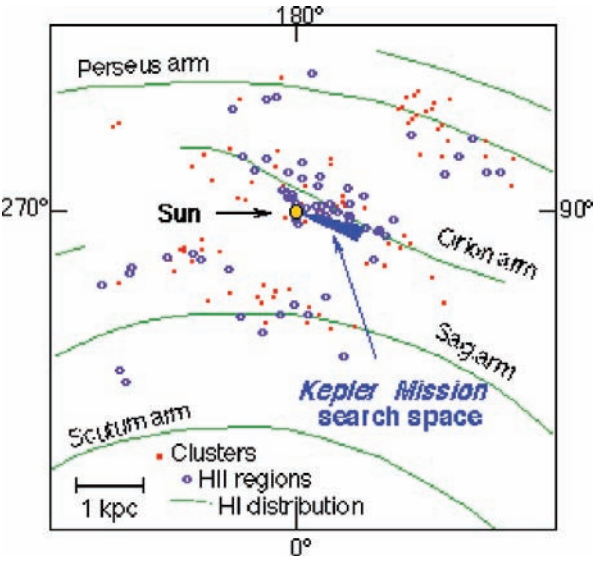

Fig. 8.1 The field that will be covered by the Kepler mission

- About 640 planets if most have a size of $R \sim 2.2\,\mathrm{R_e}$,
- About 12% with two or more planets per system.

If we also take into account that the light of the host stars can be modulated by reflected light from giant inner planets, about 870 planets with periods less than 1 week can be detected. From transits of giant planets, about 135 inner-orbit planet detections, densities for 35 inner-orbit planets, and 30 outer-orbit planet detections are predicted.

Also, in the case that no detection will be made, such a result would be extremely useful. The launch of the Kepler mission is planned for 2009.

The high precision in photometry can only be achieved by differential techniques: differential spatial photometry (the measured brightness of the target star is normalized to the average of all nearby stars) and differential temporal photometry (Transit durations are a few hours only. The brightnesses are compared to data just shortly before and after the test interval—thus, long term stability problems can be excluded.). It is also aimed that the star picture should be kept on the same pixel as long as possible to avoid possible pixel variations. The choice of the orbit of the satellite also plays a role because the scattered light background should be as small as possible.

8.3.1.3 DARWIN

The goal of the DARWIN mission will be to find Earth-like planets. The mission itself will consist of four or five free-flying spacecraft. Three spacecraft will carry 3–4 m space telescopes. From these light collectors, the light will be directed to the

central hub spacecraft. However, we must remember that the problem to observe Earth-sized planets directly is their small luminosity in the vicinity of their brilliant host star. Thus, a new technique will be applied, which is called *nulling interferometry*.

The light of the central star will be nulled by destructive interference. The light of the planet will be enhanced by constructive interference.

In pure imaging mode, DARWIN will work like a single large telescope with a diameter of up to several 100 m. Celestial objects will be observable in unknown detail. To perform the interferences mentioned above, the positions of the telescopes must be controlled with extreme accuracy (10^{-5} mm). It is planned that the system will use tiny ion engines. They will need just 5 kg of fuels sufficient for a planned duration of the mission of 5 years.

The launch date is planned for 2015.

8.3.1.4 Terrestrial Planet Finder (TPF)

The NASA TPF is scheduled for 2015. The mission will consist of multiple 3–4 m class telescopes, which must fly in precise formation. The telescopes will work in the infrared. By interferometry and interferometry nulling techniques, the planets will be detected.

The TPF has two different concepts. The visible light concept uses a single telescope with an effective diameter in the range of 6.5–8 m. This will be a coronograph, in which the star's image will be occulted to observe planets. The image contrast that is needed to be achieved in order to observe planets is $1{:}10^9$.

The infrared TPF concept should be configured as an interferometer consisting of 3–4 m class telescopes that are either spread out (over about 40 m) or operate over distances of a few 100 m on separated spacecraft. The image contrast necessary to observe planets is easier to achieve: $1{:}10^6$ in the IR.

The TPF will directly detect and survey nearby stars for planetary systems. Especially, Earth-sized planets in their HZ will be studied and by means of spectroscopy, their atmospheres should be studied. Also, biomarkers will be studied.

8.3.2 Biomarkers

Biomarkers will play an important role to answer the question whether there is life on other detected extrasolar planets. Future missions will try to focus the spectroscopic studies of extrasolar planets to detect these biomarkers.

8.3.2.1 Biomarkers on Extrasolar Planets

A biomarker is a feature whose presence or abundance requires a biological origin. Typical biomarkers are the gases O_2, O_3, H_2O, CO, and CH_4. They indicate the existence of living organisms (past or present). Also, the term *biosignature* is used.

One important biosignature is *oxygen*. If oxygen is found in an extrasolar planet's atmosphere, this would be interpreted as evidence for life. The reason is very simple: oxygen is a highly reactive gas. It will react with the minerals and metals present at the surface of a planet. Thus, a high oxygen abundance in a planetary atmosphere is not very probable unless there are lifeforms that actively produce it. These lifeforms are very familiar on Earth: the cyanobacteria and the plants. Oxygen could also be released during other processes on a planet such as during a runaway greenhouse effect (which has occurred on *Venus*).

To observe the spectral signatures of biomarkers, it is better to go to the infrared (see, e.g., Hamdouni et al., [89]) because in this wavelength region, the contrast between radiation from a host star and its planet would not be so enormous. However, oxygen does not produce strong lines in that region, but ozone (O_3) does. So it would be easier to look for ozone signatures.

But there are more sophisticate signatures for biological activity on a planet. *Plants* on Earth have a sharp increase in leaf reflectance between 700 and 750 nm. Thus, the question was studied whether such a strong increase would be detectable on extrasolar planets which are covered by a similar vegetation. This is also called the *red Earthshine* and it can be used to detect the red edge using spectroscopic observations of the dark side of the moon which is illuminated by the Earthshine. These questions have been studied by Ford, Seager, and Turne [62].

It was also suggested to use porphyrin[6] derivatives as biomarkers [241].

A completely different approach is to search for extraterrestrial biosignatures on Earth such as a search for fragments of ancient microbes that might be preserved in extraterrestrial samples (see Oehler et al. [182].

Sephton and Botta [223] gave a review on extraterrestrial organic matter and the detection of life.

8.3.3 In Situ Search for Biomarkers: Mars

There seems to exist a special Mars Curse: until 2006, only 18 of 37 launch attempts to reach Mars have been successful.

The first search for biomarkers on the Martian surface was made by the *Viking 1* and *Viking 2* landers. Viking 1 lander touched down in the Chryse Planitia and the first image was transmitted 25 seconds after landing. The lander operated for 2,245 sols until 13 November 1982. [7] Viking 2 lander touched down on the Martian surface in the Utopia Planitia on 3 September 1977. Due to a battery failure, it stopped transmission on 11 April 1980 after 1,281 sols. There were four biology experiments.

- Gas chromatograph-mass spectrometer (GCMS): vapor components can be measured and their molecular weight of each chemical can be determined. Untreated

[6] Porphyrins occur, e.g., in plants and red blood cells.

[7] The transmission broke down due to a faulty command sent by ground control.

Martian soil was partly heated partly analyzed in original state but no significant amount of organic molecules were found.

- Gas exchange (GEX): a Martian, sample was incubated with organic and inorganic nutrients. Also, water was added in a later stage. Using a gas chromatograph, it was tested whether the concentrations of gases like O, CO_2, N, H, and CH_4 changes.
- Labelled Release (LR): a sample of Martian soil was inoculated with a drop of very dilute aqueous nutrient solution (containing the Miller-Urey products).
- Pyrolytic Release (PR): in this experiment, it was tested whether some ^{14}C marked biomass can be found. An artificial ^{14}C rich Martian atmosphere was provided and if there would be some respiration process, then, part of this ^{14}C should be found in the biomass.

Other successful Mars exploration rovers are *Spirit* (landed on 3 January 2004 in the Gusev Crater, believed once to have been a crater lake) and *Opportunity* (landed on 24 January 2004 in Meridiani Planum, where there are large deposits of hematite, indicating the presence of past water).

The *Astrobiology Field Laboratory* (AFL) is an unmanned spacecraft to explore Mars and could be launched around 2016.

Summarizing the results of the biology experiments made in situ on Mars, we can state that analyzes by gas chromatography and mass spectrometry showed no organics present above the ppb level in the top few centimeters of the martian soil [18]. We have to mention, however, that the Viking mission also indicates that there is a ubiquitous layer of highly oxidizing aeolian material covering the surface. This layer could have oxidized organic material and is, therefore, responsible for the lack of detection of organic matter on the surface of Mars. This planet may, therefore, be self-sterilizing [152]. Organic degradation under simulated Martian conditions were tested by Stoker and Bullock [238]. They used chambers containing Mars-analog soil mixed with the amino acid glycine, evacuated and filled to 100 mbar pressure with a Martian atmosphere gas mixture and then irradiated with a broad spectrum Xe lamp. They found that organic compounds are destroyed on Mars at higher rates than the rate they are deposited by meteorites and no strong oxidants are needed to explain this destructive processes.

8.4 Some Case Studies of Habitability

8.4.1 Mars

As we have seen, the past conditions on Mars were more favorable for life to evolve on this planet. The pressure of the atmosphere and the average temperature were higher. There was water in liquid form. Up to 1/3 of the surface of Mars could have been covered by an ocean of about 1,700 m depth. Large amounts of CO_2 in its atmosphere led to a strong greenhouse warming, so that the surface temperature was above 0°C. The main difference in the further evolution between Earth and Mars

was because of the larger surface of the Earth plate tectonics was more active on Earth than on Mars (plate tectonics occurred only on early Mars). So, no new carbon dioxide was released in the Martian atmosphere. It is assumed that the strong greenhouse effect period on Mars lasted for 500 million years. It is certainly possible, that Martian microbes were able to adapt to the changing conditions and that these microbes have survived until now.[8]

One remarkable discovery was that of trace amounts of methane and ammonia by the Mars Express orbiter.[9] Methane is unstable in the Martian atmosphere (half-life a few 100 years), so it must be constantly replenished. There are only two ways to release methane in the atmosphere of Mars:

- Volcanism: this can be ruled out because no traces of SO_2 were found.
- Methanogens, methane producing microbes.

Onstott et al. [184] mentioned four different origins of methane on Mars: volcanogenic, sublimation of hydrate-rich ice, diffusive transport through hydrate-saturated cryosphere, and methane-generation by biological activity above the cryosphere.

The role of methane on Mars in former times to sustain a warm CO_2-rich atmosphere was discussed in the paper of Archer, Pavlov, and Smith [4]. Another source of methane could be photolysis of H_2O in the presence of CO, as it was mentioned by Bar-Nun et al. [9].

The question whether there exists life on Mars cannot be answered definitely up to now, but it is sure that the conditions on early Mars had been much more favorable for life on this planet. It is also assumed that Mars experiences ice ages like the Earth. Variations in the tilt of its rotation axis are much greater than for Earth (Sect. 4.4.2), which can trigger these ice ages as well as enhanced deposit of dark dust grains at its polar caps so that the albedo changes [83]. Fenton, Geissler, and Haberle [58] investigated documented albedo changes of large areas on the Martian surface (up to 56 million km^2), which became brighter or darker by about 10%. These variations could affect recent climate changes and larger-scale weather patterns on Mars.

8.4.2 Venus

The conditions on the surface of Venus do not seem favorable to be habitable: because of the dense atmosphere, the surface temperature ranges from 710 to 770 K. 96.5% of its atmosphere is composed of CO_2 (therefore, a strong greenhouse effect) and 3.5% is composed of N_2. The water content is 20 ppm. The mean surface pressure is 90 bar. The cloud cover starts at a height of approximately 50 km and there are droplets of sulfuric acid. Big droplets rain down from a height of approximately 50 km and, never reaching the surface, become decomposed into water, oxygen, and SO_2. Then they rise again and form sulfuric acid.

[8] See also: Research, 103, no. E12, 28529 (1998). Reference [76].

[9] The launch date was 2 June 2003.

It seems that neither the high CO_2 concentration nor the high pressure are a constraint for the evolution of life, however, the lack of liquid water and the high temperature on the surface are. In the lower and middle cloud layer of Venus, temperatures drop and there is water vapor. At an altitude of 50 km, the temperature is 40°C. The problem is osmoregulation in hygroscopic sulfuric acid clouds and the acidity in general in these clouds. Cockell [41] discusses the survival of acidophilic sulfate-reducing chemoautotrophs suspended as aerosols in the Venus atmosphere. The conditions favorable to life might have been better on early Venus. Early Venus could have been 100 times more moist than the present (this moist atmosphere of early Venus follows from an enrichment of Deuterium by a factor of 100, see Kasting [116]. For several 100 million years, on early Venus, there might have been an ocean near the boiling point, but because of the gradual increase of solar luminosity, the water in the oceans evaporated and the partial pressure of CO_2 increased because of the lack of liquid water. At that time, the atmosphere was extremely moist, however, the water vapor in the atmosphere was lost because of dissociation of H_2O in hydrogen and oxygen; the hydrogen escaped into space and the oxygen became absorbed in surface rocks.

8.4.3 Europa

Europa is one of the Galilean satellites of Jupiter. It is conceivable that life could exist there beneath a sub-surface ocean of liquid water. Moreover, undersea volcanic vents are highly probable. These are similar conditions that were found on early Earth. At the surface a temperature of the ice crust of –180°C can be expected. The surface is exposed to various hazards such as intense radiation from Jupiter's magnetosphere, impact of comets and meteoroids, and solar UV radiation. It can be estimated that below about 10 cm, these hazards become very small. At the same time, due to tidal heating (see previous chapter), the temperature increases with depth, the temperature gradient is estimated at about 0.1°C/m. Therefore, in a depth of a few km, the temperature becomes more than 0°C.

We have discussed already the threat to habitability by impacting asteroids and comets. However this has also positive effects. Impacting comets on Europa's icy crust might have provided the necessary organic compounds. This was investigated by Pierazzo and Chyba [190].

The presence of an ice ocean under the surface for other Galilean satellites was discussed already in the context of tidal heating.

8.4.4 Life on Jupiter and Other Gas Giants

The detection of extrasolar planets has clearly shown that giant Jupiter-like planets are very common in extrasolar planetary systems. Therefore, the question whether there could be life forms on such a planet is interesting. In Jupiter's atmosphere,

there are abundant biota. Ultraviolet-photoproduced [251] organic solids synthesized under simulated Jovian conditions were investigated by Khare et al. [121]. In 1976, Sagan and Salpeter [216] considered possible life forms there by comparing the conditions in Jupiter's atmosphere with terrestrial seas:

- Top level:
  - Earth: photosynthetic plankton,
  - Jupiter: sinkers, photosynthetic autotrophs.
- Middle level:
  - Earth: fish,
  - Jupiter: floaters, larger autotrophs or heterotrophs that actively maintain their pressure level.
- Lowest level:
  - Earth: marine predators which hunt fish,
  - Jupiter: hunters, organisms that seek out others.

A fourth type of Jupiter lifeforms could be the scavengers.

On Earth, life is protected by the atmosphere (from short-wavelength UV- and X-rays) and by the geomagnetic field (from energetic particles). Jupiter has both—an intense magnetic field acting as a shield and a very dense atmosphere.

8.4.5 Life in Interstellar Medium

In his novel "The Black Cloud," cosmologist Fred Hoyle described [107] a small interstellar cloud that can think and move by itself. It is very hard to imagine that such interstellar lifeforms really exist; the major drawback is that the density of interstellar matter is extremely low, so that interactions between particles are by many orders of magnitude less frequent than on planetary surfaces. Hoyle and Wickramasinghe believe that molecular clouds, such as those present in the Orion Nebula, are the most natural cradles of life [108]. Such dust clouds are composed of silicate, graphite, and iron particles. The organic molecules become polymerized on the surface of the dust grains. Extraterrestrial organic chemistry in the interstellar medium and the origins of life were discussed by Bernstein [14].

Life sustaining planets could even exist as isolated objects in interstellar space. During ordinary formation of planetary systems, Earth-sized rock and ice objects may have formed and ejected from their solar systems because of gravitational scattering from proto-giant planets. Such objects could keep H_2-rich atmospheres. Because of far-IR opacity of hydrogen, radioactive internal heat was not eliminated and these objects formed an adiabatic atmosphere where no heat loss or heat gain occurred. This leads to a paradox situation: the effective temperature of such a body will be about 30 K, however, the surface temperature could be above the melting

point of water and oceans could exist on such bodies. From the observational point of view, such objects would be extremely difficult to detect [237].

8.5 Comparison of Cosmic Catastrophes and Habitability

In this section, we compare the above mentioned candidates for life and their vulnerability to cosmic catastrophes.

8.5.1 Impacts

The main source of present asteroids that could possibly have an impact on the terrestrial planets is the asteroid main belt. On 13 April 2029, the asteroid Apophis will pass the Earth at a distance within 5–6 R_e. This 300 m asteroid will be seen with the naked eye crossing the sky at tens of degrees per hour. The theoretical impact energy would be about 1,000 megatons. It is estimated that impacts of such bodies are to be expected once every 60,000 years [6].

Let us compare the estimated impact rates for Earth and Mars for asteroids $> 1\,\mathrm{km}$:

- Earth: 0.52/Myr, rms impact velocity 11.1 km/s,
- Mars: 0.60/Myr.

Per unit area, the ratio of the present cratering rate on Mars and Earth is 3.8 (see Wetherill [265]. The age of the surface of Europa is estimated to about 10 Myr, the age of Ganymede 1 Gyr. This can be explained by a surface mixing on Europa: within a time span of 10 Myr, the surface is mixed once to a depth of 0.67 m, 10 times to a depth of 10 cm [189].

Asteroids must be perturbed on their orbits to become a potentially hazardous object. Generally, we can state that cometary-like orbits are mainly affected by Jupiter while asteroid-like orbits are mainly affected by resonances. In Table 8.2, we give the values according to Ivanov, 2006.[10]

Table 8.2 Impact rates for the terrestrial planets

Planet	Impact rate (in $10^{-15}\,\mathrm{km^{-2}y^{-1}}$)
Mercury	0.715
Venus	0.876
Earth	1.31
Mars (current orbit e=0.09	3.85
Mars (e=0.05)	2.42
Mars (e=0.01)	2.00

[10] Ivanov, 2006, http://www.lpi.usra.edu/meetings/chron2006/pdf/6013.pdf

8.5.2 Radiation Hazards

Mars does not possess a global magnetic field to shield its surface from solar flares and cosmic rays. Moreover, its atmosphere is extremely thin. Thus, there is no protection against energetic particles and radiation. Humans would have to live in shelters. It is conceivable that bacteria like *Deinococcus radiodurans* could survive such an environment. Because of these facts, many researchers place Mars still inside the HZ.

The Galilean satellites of Jupiter are exposed to energetic particles of Jupiter's intense magnetosphere. The most dangerous part of the Jovian magnetosphere is about 300,000 km above Jupiter's surface. Cooper [43] discussed habitability in high radiation environments and the special case of Europa. Radiolytic production of oxidants and simple hydrocarbons on Europa's icy surface could support evolution and survival of life within a Europan subsurface ocean. In this case, the interaction with energetic particles is essential for habitability.

8.5.3 Summary: Habitability and Cosmic Catastrophes

In Table 8.3, we give a summary of our discussion on habitability. If we assume carbon-based life, then it is evident that the best candidates would be G stars with planets in HZ. To be less strict, we can also include icy satellites near giant planets that are heated by tidal forces. Such objects are also conceivable as isolated planets and due to tidal interaction, possible satellites of such planets could be also habitable.

Cosmic catastrophes occur most frequently due to impacts. The impact rate strongly depends on the structure of the planetary system. In the solar system, there are three zones of small bodies that can be deflected by gravitational interaction of giant planets and nearby stars: the Oort cloud, the Kuiper belt, and the main asteroid belt. We can assume that similar clouds exist in extrasolar planet systems because, due to the formation of the planets, small particles are expelled to such clouds and several examples of such clouds have already been observed. Therefore, impacts are probable and could make planets inhabitable. On the other hand, they could provide important ingredients for life, such as complex hydrocarbons and water as it is assumed for the early Earth.

On the larger time scale, habitability is threatened by exploding nearby supernovae and gamma-ray burst (which may be related to supernovae).

In the previous chapter, we have also mentioned that life might exist based on different substances that are from organic chemistry; substances such as ammonia, silicon, and others may take the role of water and carbon, two fundamental elements of life on Earth. In that case, our considerations about habitability would have to be enlarged. However, such life, based on such chemistry, seems to be exotic and not very likely because of intrinsic properties of the chemical bounds themselves (unstable) and also because their occurrence is much lower than the occurrence of Carbon.

Table 8.3 Habitability and cosmic catastrophes

Object	Impacts	Radiation hazards	Galactic environment
Earth	$\sim 0.5 \times 10^{-6}\,\mathrm{yr}^{-1}$	Sun: unlikely at present	Supernovae every 10^9 yr
Mars	similar to Earth	more dangerous — no protection	similar to Earth
Europa	enhanced risk (due to Jupiter)	energetic particles from Jupiter's magnetosphere	similar to Earth
G star planet	similar to Earth	practically no hazard more dangerous,	similar in solar neighborhood
K star	similar to Earth	habitable planet closer extremely high,	similar in solar neighborhood
M star	similar to Earth	maybe no habitable planets	similar in solar neighborhood
free floating planets	no	low	dependent on location

8.6 The Drake Equation

Can we estimate the number of civilizations in our Galaxy? In 1961, *F. Drake* developed an equation to calculate that number. This equation is closely related to the Fermi paradox, which we discuss first.

8.6.1 The Fermi Paradox

We have seen that the universe is about 14 Billion years old. There was certainly enough time for the synthesis of the elements that are needed for life. It is also an observational fact that we observe many planetary systems in our neighborhood, so they must be quite common. Organic substances are found everywhere in the universe—on meteorites, in the outgassing of comets as well as in cool molecular clouds where stars are being formed.

Thus, it should be reasonable to assume that the universe, at least our galaxy, should be crowded with life. However, up to now, though intensive search programs have been made, no definite sign of extraterrestrial life has been found. Therefore, the physicist *E. Fermi* posed, in the 1950s (during lunch with his colleagues), the fundamental question, "Where are they?" [261].

There is one argument that strongly is in favor for many civilizations in the Galaxy: our galaxy contains about 200×10^9 stars. Let us assume that only the fraction 10^{-7} of these stars has planets that are habitable and have developed a higher civilization, then, there would be still 20,000 civilizations in the Galaxy. This argument states also that the Earth is not special (mediocrity principle).

The second argument is related to the first: as we know from history, the humans on Earth have always tended to spread over the whole planet—to detect new islands, continents—now, we are on the edge to explore the solar system and, maybe in some

100 years, we might even colonize some of the objects therein. Thus alien, civilizations should probably do the same and spread over the whole galaxy. There have been estimates that in a time span between 5 million to 50 million years, it would have been possible to colonize the whole galaxy. This seems an enormous time span for an individual but it is only 1/1,000–1/100 of the age of the solar system.

However, we have no indications that this really happened.

One possibility to resolve the Fermi-paradox is the so-called rare Earth hypothesis. Here, the mediocrity principle is given up. The Earth is thought an unusual or unique place at least in the Galaxy. From many philosophical and religious arguments, this hypothesis is strongly supported. It is also supported by the anthropic principle, which basically states that the only purpose of the universe is to develop human intelligence.

8.6.2 The Drake Equation

Following the discussion of *Fermi*, *Drake* (in 1960) tried to define an equation that gives the number of civilizations in our galaxy that we are able to communicate with:

$$N = R^* f_p n_e f_l f_i F_c L \tag{8.2}$$

In this equation, the following parameters enter:

R^* Rate of star formation in the Galaxy
f_p Fraction of those stars that have planets
n_e Average number of planets that are habitable per star that has planets
f_l Fraction of n_e that actually have developed life
f_i Fraction of f_l that have developed intelligent life
f_c Fraction of civilizations that have developed a technology similar to ours and that release signals
L Length of the time interval where such civilizations send out detectable signals into space (maybe also the lifetime of such a civilization).

To estimate N, Drake has used the values that are shown in Table 8.4.

The question is, how accurate can these values be determined today (see also Burchell [28])? The number of stars being formed in the Galaxy, R^*, seems well known and very close to 10/yr. n_e is unknown. Here, the new planet finding missions (GAIA, TPF, and DARWIN) will help to get better estimates. Up to now, it seems more likely to find giant gas planets near stars that are definitely nonhabitable objects. However, as it has been mentioned several times already, habitability must be extended to other objects such as satellites of gas giants (e.g., Jupiter's satellite Europa). Thus, n_e is a rather uncertain number.

The value f_l can be estimated from the Earth's history only. On Earth, life began as soon as the conditions were favorable. Abiogenesis describes the generation of life from nonliving matter, such as the generation from nonliving but self-replicating molecules near hydrothermal vents. The value of f_l would be more optimistic if life independent from Earth would be detected on Mars or on Europa.

Table 8.4 Parameters that have been used by *Drake* in his equation

Variable	Value	Comment
R^*	10/yr	10 stars formed per year in the Galaxy
f_p	0.5	half of all stars have planets
n_e	2	two planets per star are habitable
f_l	1	100% of the habitable planets form life
f_i	0.01	1% of them have formed intelligent life
f_c	0.01	1% of them are able to communicate
L	10 000 yr	lifetime of civilizations

The next two factors in Drake's equation are still more uncertain: f_i, f_C. On Earth, it lasted about 3.5 billion years from the first forms of life to the evolution of humans that are able to communicate. Thus, alien civilizations would be expected on relatively old planets.

There is an important argument against all SETI activities. If civilizations are dispersed both in space and time around a galaxy, then, the chance to interact would become extremely small. It seems, however, that times required for biological evolution on habitable planets of a galaxy are highly correlated, i.e., planets in HZ evolve nearly at the same time in a galaxy, so that at least there does not seem a dispersion in time [254].

8.6.3 Cosmic Catastrophes and Drake's Equation

The Drake-equation gives some estimate for the number of civilizations in our galaxy that are able to communicate with us at present.

However, the Drake equation only takes into account a static situation. As we have discussed, during Earth's history several mass extinctions have occurred and these are summarized in Table 8.5.[11] Several of them lead to about 90% extinction of life on Earth. We do not know exactly the reason for all of them. The questions are whether they occur on other habitable planets as well and whether they lead to total extinction, etc. Current theories suggest that these catastrophes to habitability on Earth were caused by climate changes triggered by large asteroid impacts.

Whatever the cause of such mass extinctions, the simple Drake equation has to be modified [39]. In order to evolve from primitive to higher forms of life, the habitability of a planet must not be disturbed for about 3.5 billion years.

In the previous chapters, many possible cosmic catastrophes that can destroy the habitability on a planet were described:

- Variations of the central host star—large flaring events or CME like events can considerably influence on the atmosphere of a habitable planet.
- Impacts: large impacts lead at least to mass extinctions; this is a well known fact for evolution of life on Earth.

[11] Adapted from: Ref. [48].

Table 8.5 Some known mass extinction in Earth's history. Percentage means the percent of species extinct

Period	Time (BP)	Percentage
Ordovician	444 mill.	85
Devonian	370 mill.	83
Permian	250 mill.	95
Triassic	210 mill.	80
Cretaceous	65 mill.	76

- Changing parameters of a planet's orbit: the presence of giant planets inside and possibly outside a HZ around an extrasolar planetary system leads to great perturbations of the orbital parameters of planets in a HZ.
- Variations of the obliquity: the obliquity of the rotation axis of a planet strongly influences on the severity of seasons and if there occur large variations during the long time interval needed for life to evolve to higher forms, life may have become extinct several times, thus, no chance for, other than perhaps microbial, life to evolve. In the case of the Earth, the Moon has a stabilizing effect.
- Galactic environment: we have discussed a HZ around a galaxy. However, the environment changes as a planetary system orbits the center of a galaxy.
- Shielding effects: the Earth lies within the heliosphere that is due to the extended solar magnetosphere.
- Presence of giant gas planets outside the HZ: it has been shown that due to the strong gravitational attraction of Jupiter many cometary like bodies are deflected and, thus, the threat of a collision with planets in the HZ becomes considerably smaller.

Thus, to find a planet similar to our present, the following conditions must be fulfilled: $f_\odot$ denotes the number of stars that are solar like (spectral type G) and are at least 4 billion years old and in the HZ of the Galaxy. N_{pl} is the fraction of these stars that have Earth-sized planets. N_M is the fraction of this Earth-sized planets that have a large satellite that stabilizes their rotational axis or where there is so little variation in the obliquity of their axis that no large climate variations occur. Let us estimate these numbers.

Of all the main sequence stars, about 8% are of type G. If we assume that the Galaxy contains 200×10^9 stars, then, there would be 16 Billion G type stars. Let us further assume that 10% of these stars are in the HZ of the Galaxy, then, there remain 1.6 Billion stars. The occurrence of planetary systems seems quite common, so we can assume that 1/3 of these stars has planetary systems, thus, about 0.5 billion planetary systems in the HZ of the galaxy remain. Since the HZ around G stars is relatively broad, the probability to find planets there is quite high, however, there should be an Earth-sized planet with a large satellite that stabilizes its rotational axis. If the percentage of G stars that have planetary systems and a planet of Earth size in its HZ is 0.1% then, there still remain 500,000 objects.

Thus, our crude estimate gives 500,000 Earth-like planets in the HZ of the Galaxy that are around a G-type star and that lie in the HZ with stable axis of rotation. Even

if we assume that from these 500,000 objects only 1/1,000 contain a continuous habitable zone (no nearby supernova explosion, no catastrophic impacts...), still 500 object remain.

However, we must take into account the age of these 500 objects. As we know from the Earth, it took about 3.5 billion years for life to develop to higher forms. How many of these 500 objects are old enough? If we take a crude value of 1/10, then, finally, our very crude approximation gives 50 living civilizations in the galaxy in the HZ which are roughly at the same level than we are and which have survived cosmic catastrophes resp. no major catastrophes have occurred on these planets. Taking into account the enormous distances in the galaxies, we immediately see how difficult communications with these civilizations would be. This estimation could explain why they did not answer yet.

Finally, maybe the above made estimation is even too pessimistic. We have considered only G stars. Maybe even the cooler K stars could be candidates for hosting planets in HZ. The percentage of these stars is 13% of all main sequence stars. Civilizations orbiting a K-star could profit from a longer main sequence lifetime of their host star. If life would have been extinct there by one of the cosmic catastrophes described above and in the previous sections, then maybe there was a "second chance" or there will be one.

Taking K stars with the argument of a "second chance" for life into account this would add another say 100 civilizations to our galaxy.

Finally, as it was discussed, some of the large satellites orbiting giant planets and maybe even isolated free floating planets are possibly habitable. Thus, the number of habitable worlds strongly increases by some unknown factor.

While all these estimates are very crude but also very conservative, we still may expect at least 100–200 civilizations that are similar to our civilization in our galaxy. This is not a large number but definitely different from one, because at least to my opinion, the parameters to calculate the values have been taken quite conservatively. Let us hope to find them!

Chapter 9
Appendix

9.1 Life and Chemistry

In this section, the chemistry of life is outlined. Organic chemistry, the study of compounds in which carbon is the principal element, describes the complex molecular reactions that are found in biological systems. We assume that life can only evolve on the basis of carbon compounds; other alternatives were discussed previously.

9.1.1 Atoms and Bonds

The smallest unit of a chemical element is the atom. Atoms consist of a nucleus, which contains positively charged protons *p*, neutral neutrons *n*, and electrons e^-, which are located in discrete shells (orbitals) around the nucleus. Atoms are characterized by atomic number, which is the number of protons, and the atomic mass, which is the number of protons plus neutrons. The mass of an electron is only $\sim 1/1,840$ that of a proton, so their contribution to the atomic mass is negligible. The element nitrogen with atomic number 7 and atomic mass 14 is written as $^{14}_{7}\mathrm{N}$ (so the nucleus contains seven protons and seven neutrons). The chemical elements are defined by the atomic number. Isotopes of an element contain the same number of protons but the number of neutrons is different. For example, one isotope of oxygen contains 8 p and 8 n, another one 8 p and 10 n.

Molecules consist of two or more atoms bound together, examples are H_2 or the water molecule H_2O.

The number of electrons in the outermost orbital determines how many chemical bonds can be formed. There are three different types of chemical bonds:

1. Covalent bond: a pair of electrons is shared in the outermost orbital. Consider H_2: the two H-atoms share their single electrons making a combined orbital. This is written as

$$\mathrm{H}--\mathrm{H} \tag{9.1}$$

   For life, carbon chemistry is especially important. This element has electrons in two orbitals: two in the innermost and four in the outer orbital, which can be

A. Hanslmeier, *Habitability and Cosmic Catastrophes*, Advances in Astrobiology and Biogeophysics, DOI 10.1007/978-3-540-76945-3_9,

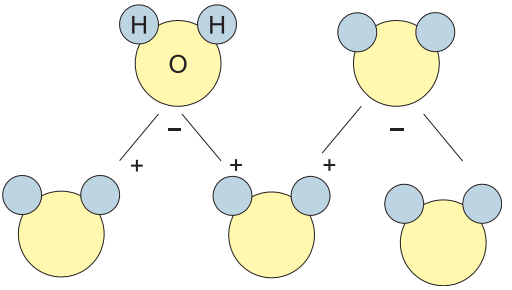

Fig. 9.1 Cohesion of water molecules. The H-atoms are slightly positively charged, the O-atoms negatively

shared. Thus, carbon combines with, e.g., four H-atoms. Polar covalent bonds are formed when electrons are closer to one atom than to another — the molecule becomes polar. Water, another basic substance for life, is a well-known example of a polar molecule (Fig. 9.1). The shared electrons are more attracted to the oxygen atom, so that the H-atoms become slightly positive and the O-atoms negative.

2. Ionic bonds: molecules that lose or gain electrons are called ions. In sodium chloride (table salt), NaCl, the Na-atom becomes positively charged, and the chlorine-atom negatively charged; Na loses one electron in the outermost orbital, Cl gets one electron:

$$Na^{+}Cl^{-} \tag{9.2}$$

   NaCl plays in important role for the chemistry of biological systems.
3. Hydrogen bonds: they form as a result of attraction between positively charged H-atoms in polar molecules and negatively charged atoms in other polar molecules. This bond is weak (only 7–10% of the strength of covalent bonds). For biochemical processes, the H-bonds are extremely important because they maintain the shapes of proteins, etc.

Generally, valence is the combining capacity of atoms or ions. Valences must be balanced in any chemical reaction. Chlorine has a valence of -1, calcium (Ca), which is important in plant cell walls, has a valence of +2. Hence, the chemical formula for calcium chloride has to be $CaCl_2$ and we write also Ca^{++}, Cl^{-}.

9.1.2 Acids, Bases, and Salts

As we discussed before, water molecules are held together by weak H-bonds. In pure water, a few molecules dissociate into H^{+} and hydroxyl ions OH^{-} and an equilibrium forms. Acids are chemicals that release H^{+} ions when dissolved in water. There are more H^{+} ions than hydroxyl ions in the solution. Weak acids release only few ions, whereas strong acids, like sulfuric acid, dissociate almost completely

into H^+ and sulfate ions. Bases (alkaline compounds) release negatively charged hydroxyl ions (OH^-) when dissolved in water. Sodium hydroxide (NaOH) dissociates in water into Na^+ and OH^-. The acidity or alkalinity of the fluids of cells has to be stable for chemical reactions vital to life to occur.

The concentration of H^+ ions defines the pH-scale. Water has a pH of 7, there is an equal number of H^+ and OH^- ions. The lower the pH-value the higher the degree of acidity. To name two organic substances, vinegar has $pH = 3$ and egg has $pH = 8$. Precipitation with $pH \leq 4.5$ is referred to as acid rain.

When mixing an acid with a base, the H^+ ions of the acid bond with the OH^- ions of the base forming water. The remaining ions bond together forming salt, like in the reaction:

$$HCl + NaOH \rightarrow H_2O + NaCl \tag{9.3}$$

9.1.3 Important Elements for Life

In Table 9.1, some important elements for life are given. The reactions where these elements are involved in will be discussed in the section about biochemical processes.

By mass, human cells consist of 65–90% water and 99% of the mass of the human body is made up of the six elements, which are oxygen (65%), carbon (18%), hydrogen (10%), nitrogen (3%), calcium (1.5%), and phosphorus (1.2%).

Table 9.1 Some important elements for life. *A* denotes atomic number, *M* atomic mass

Element	A	M	Functions
Hydrogen (H)	1	1	Part of nearly all organic molecules
Carbon (C)	6	12	Skeleton of organic molecules
Nitrogen (N)	7	14	Amino acids, nucleic acids, chlorophyll
Oxygen (O)	8	16	Respiration, part of most organic molecules
Magnesium (Mg)	12	24	Chlorophyll (basic element)
Phosphorus (P)	15	31	Part of ATP, nucleic acids
Sulfur (S)	16	32	Stabilizes protein's 3D structure
Potassium (K)	19	39	Balance between ions in cells
Calcium (Ca)	20	40	Structure of cell walls
Iron (Fe)	26	56	Oxygen transport during respiration

9.1.4 The Element Carbon

The name comes from the Latin word *carbo* which means coal. The nonmetal element Carbon C has an atomic number of 6 and an atomic mass of 12.0107 amu.[1]

[1] 1 amu (atomic mass unit) = $1.66053886 \times 10^{-27}$ kg.

Table 9.2 Isotopes of carbon

Isotope	Half life
^{11}C	20.3 min
^{12}C	stable
^{13}C	stable
^{14}C	5730.0 years
^{15}C	2.5 sec

The melting point is 3,500°C and the boiling point 4,927°C. At a temperature of 20°C, carbon has a density of 2.62 $g\,cm^{-3}$. The different isotopes of carbon are listed in Table 9.2. The outer valence shell of Carbon contains only four electrons but could contain up to eight electrons, therefore, Carbon can share electrons with up to four different atoms and this explains why carbon can form extremely complex molecules that are found in amino acids or proteins.

There are several known allotropes[2] of carbon: amorphous, graphite, diamond, and fullerene.[3] Graphite, a very soft and slippery substance, and diamond, the hardest substance we know, are both made of carbon atoms; their differences result from the arrangement of the atoms: in graphite, the carbon atoms form many layers, in each layer there are strong covalent bonds but the forces between the different layers are very weak. In diamond, each carbon atom is the same distance to each of its neighboring atoms, thus, a rigid network of C-atoms is constructed (Fig. 9.2). The fullerenes were discovered by R. Buckminster Fuller in 1985 (Fig. 9.3).

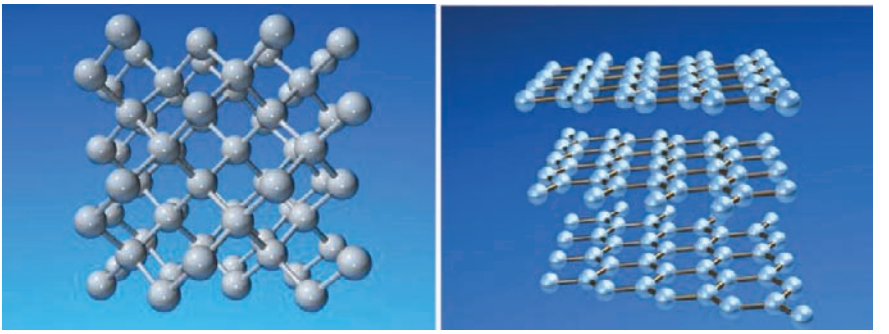

Fig. 9.2 Two allotropes of carbon: diamond (*left*) and graphite (*right*)

[2] An allotrope means a different form of an element.

[3] In fullerenes, the atoms are ordered on a football like structure.

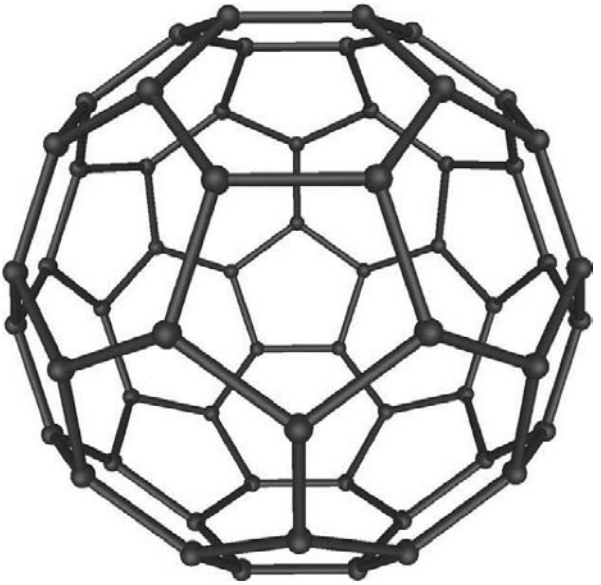

Fig. 9.3 Example of a buckyball fullerene (after M. Ströck)

9.1.5 Hydrocarbons

Hydrocarbons consist only of carbon and hydrogen. There are four groups of hydrocarbons.[4] They are (i) Alkanes (paraffin series): single covalent bonds between the carbon atoms, designation: "...ane"; (ii) Alkenes: with a double covalent carbon to carbon bond, designation: "...ene"; (iii) Alkynes: carbon to carbon triple bond, "...yne" and (iv) Aromatic hydrocarbons. The first three groups consist of straight or branched chains of carbon atoms, whereas aromatic hydrocarbons have a ring-like structure. A six-carbon ring structure has three double bonds which are not localized but spread over the whole molecule. Such a C_6H_6 ring structure, called benzene, is more stable. For example, in phenol, one H at the top is replaced by an OH.

The simplest hydrocarbon is *methane* (CH_4) but it is more instructive to write down the structural formula (Fig. 9.4). The next alkane consisting of two C-atoms is

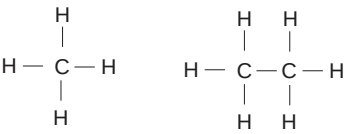

Fig. 9.4 Methane, ethane

[4] Nomenclature by International Union of Pure and Applied Chemistry (IUPAC).

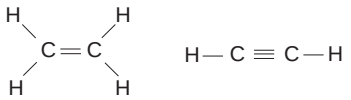

Fig. 9.5 Ethene, ethyne

ethane (C_2H_6). The unsaturated *ethene* (Fig. 9.5) has a double bound between the carbon atoms: C_2H_4; the *ethyne* has a triple bond: C_2H_2.

The general formula for the alkanes is C_nH_{2n+2}, for the alkenes C_nH_{2n}, and for the alkynes C_nH_{2n-2}.

The following nomenclature is for the alkanes. Methane: CH_4, Ethane: C_2H_6, Propane: C_3H_8, Butane: C_4H_{10}, Pentane: C_5H_{12}, Hexane, Heptane, Octane... For the alkenes, the simplest compound is ethene C_2H_4, then butene C_3H_6... and for the alkynes the simplest compound is ethyne C_2H_2, etc.

If the arrangement is in a straight chain, then we use the term *normal* (abbreviated as *n*). Compounds with the same molecular formulae but with different structures are called *isomers*, *iso* becomes the prefix. An example is *n*-butane (boiling point -0.5^o) and isobutane (boiling point -10^o).

Unsaturated compounds do not contain the maximum number of hydrogen atoms. For example, ethene (also called ethylene) can add more hydrogen atoms breaking the double bonds to form ethane:

$$C_2H_4 + H_2 \rightarrow C_2H_6 \tag{9.4}$$

Petroleum is a mixture of alkanes, cycloalkanes,[5] and some aromatic hydrocarbons. As a rule, we can state that the larger the molecule, the higher the boiling point (so the products may be obtained from distilled vapors). Methane, Ethane occur in *natural gas*, pentane, hexane up to dodecane in *gasoline*, tridecane up to pentadecane in *kerosene*, and hexadecane up to octadecane in *diesel fuel*.

The hydrocarbons account for 5% of the known organic compounds. However, in 95 % of all organic compounds, can be considered hydrocarbon derivatives, one or more H atoms have been replaced by some element or a group of elements. Two examples of organic halides are *Chloroform* ($CHCl_3$) and *Vinyl chloride* (C_2H_3Cl), where Cl is replaced by H.

One final important property of alkene molecules is that they can add to each other to form a very long chain (up to several hundreds of molecules): when Ethene (ethylene) is heated under pressure with a catalyst, the heating energy breaks up the double bonds, and single covalent bonds occur allowing long chains to be formed—*polyethylene*. This process is called *polymerization*.

Before discussing further groups of compounds, we make a generalization:

- Functional group *R*: the site of the chemical reaction defines the chemical properties of an organic compound.
- R': stands for one or more hydrocarbon groups.

[5] These are types of alkanes that have one or more rings of carbon atoms in the chemical structure of their molecules.

Methane and ethane can be found on different bodies of the solar system including the gas planets Jupiter, Saturn, Uranus, and Neptune as well as in the tails of comets. On Earth, the methane concentration in the atmosphere is 0.0001 %. It is a very strong greenhouse gas. *Methanhydrate* is methane frozen in ice on the deep ocean grounds and can be used as fuel in the future.

9.1.6 Alcohols and Organic Acids

An alcohol is an organic compound formed by replacing one or more H on an alkane with a hydroxyl functional group, OH.

The simplest alcohol is methyl alcohol (*methanol*) (Fig. 9.6): CH_3OH. *Ethanol* is C_2H_5OH and it is produced by *yeast* acting on sugars:

$$\boxed{\text{Sugar + action of yeasts} \rightarrow \text{ethanol} + CO_2.}$$

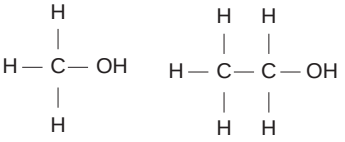

Fig. 9.6 Methanol, ethanol

An alcohol with two OH groups is called *glycol*. Ethylene glycol, $C_2H_4(OH)_2$ is used as an antifreeze. An alcohol with three OH groups is called a *glycerol*, $C_3H_5(OH)_3$.

The term *organic acid* comes from acids that were derived from organisms (e.g., distillation of ants gives formic acid). The general formula is RCOOH.

Formic acid: HCOOH also causes the sting of bees, ants, and plants (stinging nettle), see Fig. 9.7 left. *Acetic acid* (vinegar): formed from oxidation of ethanol (Fig. 9.7, right): $C_2H_5OH \rightarrow CH_3COOH$.

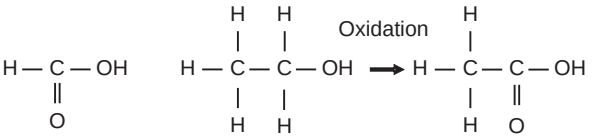

Fig. 9.7 Formic acid (*left*), acetic acid (acid of vinegar) is formed by oxidation of ethanol (*right*)

9.1.7 Organic Compounds of Life

Organic molecules in living things are macromolecules with molecular weights of thousands or millions of atomic mass units. These macromolecules are comprised of repeated monomers. The reaction to combine monomers to polymers is

a dehydration synthesis. The reverse of this reaction is hydrolysis: by addition of water larger polymers are split, e.g., digestion of food molecules in the stomach.

There are three types of polymeric macromolecules: carbohydrates, proteins, and nucleic acids. The lipids are not true polymers but macromolecules.

9.1.7.1 Carbohydrates

They are composed of carbon, hydrogen, and water; the name means "watered carbon". For each carbon atom, there is one water molecule, thus, the number of carbons and oxygens is the same and the number of hydrogens is the double of that. Examples are $C_3H_6O_3$ and $C_5H_{10}O_5$. These are simple *sugars* and their ending is *-ose*. The above examples are *triose* and *pentose*.

For living organisms, the carbohydrates serve as immediate source of energy (sugar), provide shape to cells (*cellulose* in plant cell walls), and are an essential part of genes (DNA). Simple sugars can be combined: the combination of two sugars results in a *disaccharide*, an example is the formation of *sucrose* (ordinary table sugar) when glucose and fructose are joined with the loss of water (Fig. 9.8).

$$C_6H_{12}O_6 + C_6H_{12}O_6 \rightarrow C_{12}H_{22}O_{11} + H_2O \tag{9.5}$$

Other common disaccharides are *lactose* (milk sugar) and *maltose* (malt sugar).

Examples of polysaccharides are cellulose, starch, and glycogen. *Cellulose* is used in the plant cell walls and cannot be digested by humans (only by ruminants, e.g., cows, sheep and termites—they have microorganisms within their digestive tracts that digest cellulose). Thus, plant cell walls add fiber to our diet. *Starch* is digestible by most organisms and *glycogen* is found in muscle cells.

Sugar molecules are also part of *DNA* (deoxyribonucleic acid), RNA (ribonucleic acid), and *ATP* (adenosine triphosphate).

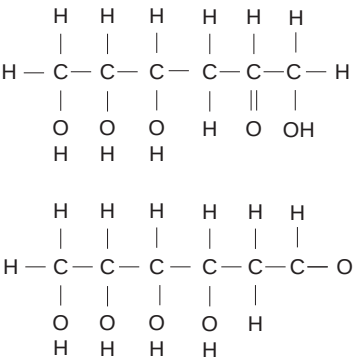

Fig. 9.8 Fructose (*above*) and glucose (*below*)

9.1.7.2 Proteins

Proteins are polymers made up of monomers—the *amino acids*. An amino acid (Fig. 9.9) contains a short carbon skeleton, a carboxylic acid group, an amino group (a N and two H), and side chains.

Any amino acid can form a bond with any other amino acid. There are 3 million possible combinations for a molecule that is only five amino acids long. The combinations are called *polypeptide chains*.

There are four levels of protein structure. The sequence of amino acids in a polypeptide structure is controlled by the genetic information of the organism. The *genes* tell the cell to link particular amino acids in a specific order—they determine the polypeptide's primary structure. The string of amino acids can twist—the twisted parts are called secondary structure. The tertiary structure determines different portions of the molecule—the complex three-dimensional structure. The quaternary structure results from combinations of tertiary structures. Such structures can be found in antibodies (*immunoglobulins*) and hemoglobin molecules. If energy is added to the system, the bonds within the protein molecules may break, and the protein becomes denatured. A well-known example is when the gelatinous structure of an egg is formed into a white solid when cooking it.

Proteins play different roles: *hemoglobin* serves as carrier of oxygen[6] in red blood cells (erythrocites), *collagen* provides support and shape, chemical messengers (*hormones*), and *enzymes*—they act as catalysts.

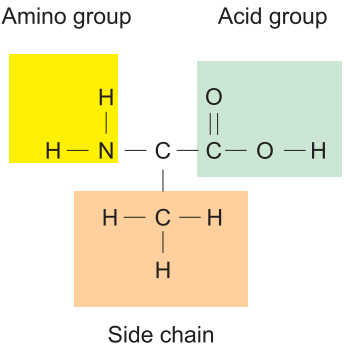

Fig. 9.9 Amino acid

9.1.7.3 Nucleic Acids

All nucleic acids are complex polymeric molecules constructed from monomers, also called *nucleotides*. These nucleotides consist of

[6] Here, Fe plays an important role; deoxyhemoglobin is the form of hemoglobin without the bound oxygen. The oxyhemoglobin has significantly lower absorption of the 660 nm wavelength than deoxyhemoglobin, while at 940 nm, its absorption is slightly higher. This difference is used for measurement of the amount of oxygen in a patient's blood by an instrument called pulse oximeter.

1. 5-carbon single sugar molecule group:
   - Ribose, $C_5H_{10}O_5$
   - Deoxyribose, $C_5H_{10}O_4$
2. Phosphate group
3. Nitrogenous base—double-ring: Adenine, Guanine; single-ring: Thymine, Cytosine, Uracil.

The monomers form polymers that combine to make large polymers and sugar and phosphate form the backbone—the nitrogenous base at the other side.

The deoxyribonucleic acid (DNA) (Fig. 9.10) is composed of two strands forming a double-ladder. The two strands are attached by the base pairs: Adenine–Thymine (A–T) and Guanine–Cytosine (G–C).

The coding strand is the basis of the genetic code and consists of sequences like TAG or ACT or CAT. The opposite strand is called noncoding. It protects the coding strand from chemical and physical damage. The strands are twisted into a helix containing tens or thousands of bases pairs (AT and GC). This can be read like chapters—each gene is a chapter. Different genes are defined by beginning and ending statements. The human body cells contain 46 strands of DNA, each is composed of thousands of genes (chapters). The strands are called *chromosomes*.

One cell contains 2 meters of DNA! All chromosomes unravelled from the cells of a human body would make 6,000 times the Earth-Moon distance. The DNA can be found in the nucleus of Eukaryotic cells and it is free floating in Prokaryotic cells.

In the ribonucleic acid (RNA), thymine is replaced by uracil and, therefore, adenine pairs with uracil.

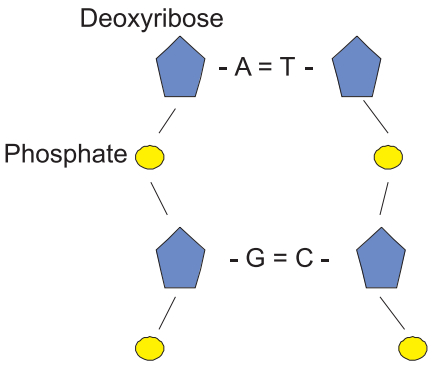

Fig. 9.10 Double strand of DNA. The backbone is sugar (deoxyribose) and phosphate

9.1.7.4 Lipids

There are three different types of lipids:

true fats (pork fat or olive oil), phospholipids (cell membrane component), steroids (most hormones).

True (neutral) fats are complex organic molecules that provide energy.
1 g of sugar provides 4 calories,[7]
1 g of fat provides 9 calories.

The building blocks of a fat are a glycerol molecule and fatty acids. A *glycerol* is a carbon skeleton that has three OH groups attached to it. The chemical formula for glycerol (structural formula Fig. 9.11) is

$$C_3H_5(OH)_3 \tag{9.6}$$

Glycerol is used as glycerin (the IUPAC designation is 1,2,3-trihydroxypropane) as an additive to many cosmetics to make them smooth.

The *fatty acids* in fats have a long-chain carbon skeleton consisting of a carboxylic acid functional group and can be saturated (as many hydrogen bonded as possible) and unsaturated. The saturated fats are found in animal tissues and tend to be solids at room temperature (e.g., butter). Unsaturated fats are frequently plant fats or oils—they are usually liquids at room temperature. Besides storing energy, fats in animals can also provide a protection layer (e.g., heat loss reduction for whales).

A fat that is formed from a glycerol molecule and three attached fatty acids is called a *triglyceride*; this is 95% of the fat stored in human tissue.

Phospholipids are complex water-insoluble organic molecules—they resemble fats but contain a phosphate group (PO_4). The phospholipids are a major component in membrane cells and they separate the membranes from the exterior environment.

The steroids contain an arrangement of interlocking rings of carbon. They often serve as hormones helping to control several processes in the body. One example is *cholesterol*.[8] Cholesterol regulates membrane fluidity.

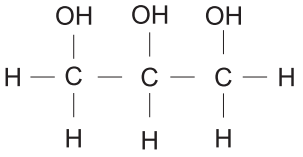

Fig. 9.11 Glycerol

9.2 Stars and Radiation

9.2.1 Electromagnetic Radiation

The spectrum of electromagnetic radiation ranges from gamma rays , X-rays, UV, visible light to IR, microwaves, and radiowaves. This is also a sequence of

[7] 1 cal = 4.1868 J.

[8] Most cholesterol is synthesized in the body, some is of dietary origin. LDL (low density lipoprotein) is harmful; HDL (high density lipoprotein) is thought to be beneficial.

decreasing energy and increasing wavelength (decreasing frequency) of the radiation. The wavelength λ and frequency ν are related by

$$c = \lambda \nu \tag{9.7}$$

where c=300,000 km/s, the speed of light. The energy of electromagnetic radiation (Fig. 9.12) depends on its frequency:

$$E = h\nu \tag{9.8}$$

h is the Planck constant, $h = 6.626 \times 10^{-34}\,\mathrm{J \cdot s}$.

Some examples of energy of different types of radiation:

- Visible light: $\sim$ eV,
- X-rays: $\sim 10^2$eV,
- Gamma rays: $\sim 1,000 \times 10^9$eV.

To a good approximation, we can assume that stars radiate like a black body. A black body is a theoretical idealization: an object that absorbs completely all radiation at all wavelengths. The radiation of a black body at temperature T_S is given by the *Planck law*:

$$I_\nu = B_\nu = (2h\nu^3/c^2)/\exp(h\nu/kT_S) - 1 \tag{9.9}$$

Here, I_ν is the intensity of radiation at frequency ν; h, k, and c are Planck's constant, Boltzmann's constant, and the speed of light, respectively. $k = 1.38 \times 10^{-23}\,\mathrm{JK}^{-1}$.

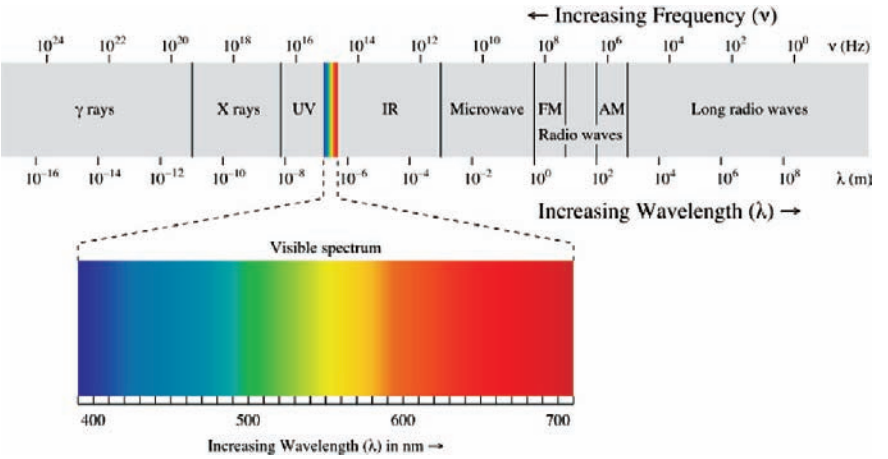

Fig. 9.12 Electromagnetic radiation. Only the visible part (400–700 nm), some part of the IR, and the radio part is observable from the ground. Wikimedia Cosmos

9.2.2 Spectral Lines

By observing spectral lines, we can deduce many important parameters: the elements that emit these lines, relative velocities between source of emission and observer (Doppler effect), temperature, pressure, etc.

Spectral lines can be in emission or absorption. In atoms, the electrons are found on defined orbital states. To make a transition to a higher state, energy in the form of a photon is required—this causes an absorption line; when the electron jumps back to its original state (which depends on the temperature and pressure), then, energy is emitted in form of an emission line. Electrons can only be found in defined energy states (see textbooks on quantum mechanics).

Let us calculate a very simple example: what is the energy released when in a hydrogen atom, a transition from orbital states n=3 to n=2 occurs? The formula is (R_H is the Rydberg constant)

$$\frac{1}{\lambda} = \left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) \times R_H \qquad R_H = 1.0973731568 \times 10^7\,\mathrm{m}^{-1} \tag{9.10}$$

For the wavelength, $\lambda = 6,563nm$ is obtained, which corresponds to 1.9 eV. The energy from $n = 1$ to ∞ corresponds to 91.13 nm or 13.6 eV. This energy is needed to ionize a hydrogen atom from its ground state (Lyman series). The lines of the Balmer series are in the visible part of the spectrum and denoted as Hα ($n = 3$ to $n = 2$), Hβ ($n = 4$ to $n = 2$), etc. and they converge to $n = 2 \rightarrow \infty$ and $\lambda = 364.51$ nm.

The electron volt, eV, is given by:

$$1\,\mathrm{eV} = 1.602 \times 10^{-19}\,\mathrm{J} \tag{9.11}$$

Energy, frequency, and wavelength of a photon are related by

$$E = h\nu = \frac{hc}{\lambda} = \frac{1240\,\mathrm{nm}}{\lambda}\,\mathrm{eV} \tag{9.12}$$

9.2.3 Stellar Parameters

Stars are gaseous spheres being most of their lifetime in hydrostatic equilibrium. The only information we can directly obtain from a star is its radiation and position.

To understand the physics of stellar structure, stellar birth, and evolution, we have to derive quantities such as stellar radii, stellar masses, composition, rotation, magnetic fields, etc. We will just very briefly discuss how these parameters can be obtained for stars.

- Stellar distances: a fundamental but not an intrinsic parameter. Stellar distances can be measured by determining their annual parallax, that is, the angle the Earth's orbit would have seen from a star. This defines the astrophysical distance

unit *parsec*. A star is at a distance of 1 parsec (pc) if its parallax is $1''$. 1 pc = 3.26 light years.[9]

- Stellar temperatures: can be derived from Planck's law (see Sect. 9.2.1).
- Spectral lines indicate which elements are present. By measuring the wavelength λ of an observed stellar spectral line and comparing it with the laboratory wavelength λ_0, we can determine the radial velocity v_r of the star:

$$\frac{v_r}{c} = \frac{\lambda - \lambda_0}{\lambda_0} \tag{9.13}$$

  If the lines are blueshifted, then, the object approaches the observer; if they are redshifted, the objects moves away. If the star expands, then, the lines will also be shifted.
- Stellar radii: once the apparent diameter of a star is known, its real diameter follows from its distance d. The problem is to measure apparent stellar diameters, which are extremely small because of the large stellar distances. One method is to use interferometers and another method is to use occultation of stars by the moon or mutual occultations of stars in eclipsing binary systems. Consult general textbooks about astronomy.
- Stellar masses: can be determined by using Kepler's third law in case we observe a binary system. Stellar evolution strongly depends on stellar mass, however, we know accurate masses only for some 100 stars.
- Once mass and radius are known, the density and the gravitational acceleration follow. These parameters are important for the stellar structure.
- Stellar rotation: for simplicity, we can assume that a star consists of two halves: one half approaches to the observer and the spectral lines from that region are blueshifted, the other half moves away and the spectral lines from that area are redshifted. The line profile we observe in a spectrum is a superposition of all these blue- and redshifted profiles, therefore, stellar rotation causes a broadening of spectral lines;
- Stellar magnetic fields: magnetically sensitive spectral lines are split into several components under the presence of strong magnetic fields. This is called the *Zeeman effect* and the amount of splitting $\Delta\lambda$ depends on

$$\Delta\lambda \sim \lambda^2 H g \tag{9.14}$$

  g... Landé factor, follows from quantum mechanics. Lines with a $g = 0$ are not split. H... magnetic field strength. Note that the amount of splitting depends on the square of the wavelength, so that this effect is larger in the IR.

[9] 1 Lyr is the distance that light propagates in 1 year at a speed of 300,000 km/s. The distance is thus 365 $\times$ number of seconds per day (86,400) $\times$ speed of light (300,000 km/s) = 10^{13} km.

9.2.4 Stellar Spectra, the Hertzsprung-Russell Diagram

The analysis of stellar radiation is fundamental for the derivation of physical quantities describing a star. Putting a prism or a grating inside or in front of a telescope, we obtain a spectrum of a star. Such a spectrum contains many lines, most of them are dark absorption lines. Each chemical element has a characteristic spectrum.

In the HRD, the temperature of stars is plotted versus brightness (Fig. 9.13). The temperature of a star is related to its color (Wien's law): blue stars are hotter than red stars. In the HRD, the hottest stars are on the left side. The temperature increases from right to left. Stellar brightness is given in *magnitudes*. The magnitude scale of stars was chosen such that a difference of 5 magnitudes corresponds to a factor of a 100 in brightness. The relation between intensity and magnitude is given by

$$\text{Magnitude} = \text{const.} - 2.5 \log(\text{Intensity}) \tag{9.15}$$

The smaller the number (which can be even negative) the brighter the star. The brightest planet Venus, e.g., has magnitude $-4.^{\text{m}}5$ and the Sun has $-26.^{\text{m}}5$. The faintest stars that are visible to the naked eye have magnitude $+6.^{\text{m}}0$. Since

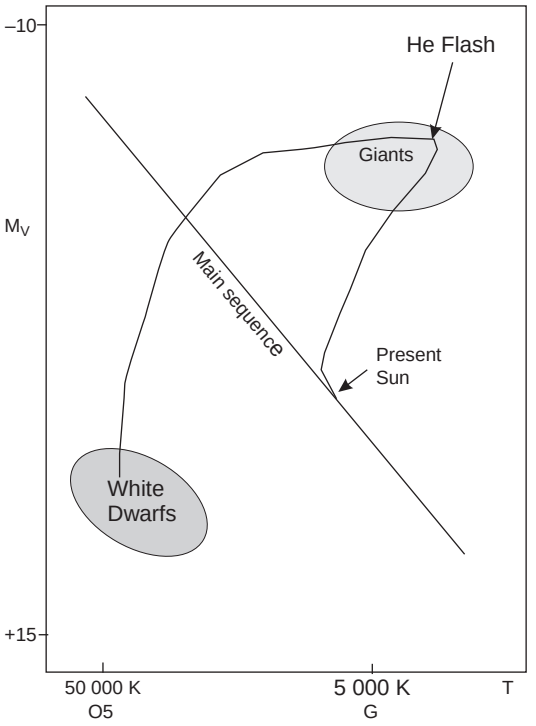

Fig. 9.13 Sketch of the Hertzsprung-Russell diagram with evolutionary path of the Sun

apparent magnitudes depend on the intrinsic luminosity and the distance of a star, *absolute magnitudes* (designated by $^{\mathrm{M}}$) are defined as the magnitude a star would have at a distance of 10 pc. In the HRD, often absolute magnitudes are plotted as ordinates instead of luminosity. The relation between m and M is given by

$$m - M = 5 \log r - 5 \tag{9.16}$$

r is the distance of the object in pc. The Sun has $M = +4.^{\mathrm{M}}5$; seen from a distance of 10 pc, it would be among the fainter stars but still visible with the naked eye.

How can we determine stellar temperatures? If the Planck equation (9.9) is integrated over all frequencies (wavelengths), we obtain a formula for the total power emitted by a black body, the *Stefan-Boltzmann* law:

$$\int_0^\infty B_\lambda d\lambda = \sigma T^4, \tag{9.17}$$

and for the luminosity of a star:

$$L = 4\pi r^2 \sigma T_{\mathrm{eff}}^4 \tag{9.18}$$

For the Sun, $T_{\mathrm{eff}} = 5,785\,\mathrm{K}$. This formula defines the effective temperature T_{eff} of a star. $\sigma = 5.67 \times 10^{-8}\,\mathrm{Wm^{-2}K^{-4}}$ is the Stefan-Boltzmann constant.

What is the power emitted per unit area on the Sun's surface? Answer: put T = 6,000 K, we find that the Sun radiates 70 MW per m^2 of its surface.[10]

By taking the derivative with respect to λ of Planck's law and setting it equal to zero, one can find the peak wavelength, for which the intensity is at maximum:

$$T \lambda_{\max} = 2.9 \times 10^{-3}\,\mathrm{mK} \tag{9.19}$$

This is also called *Wien's law*.

At about which wavelength can planets be expected to radiate most of their energy? Answer: Let us assume the temperature of the Earth = 300 K. Then,

$$\lambda_{\max} = 2.9 \times 10^{-3}/300 \sim 10\,\mu \tag{9.20}$$

The Sun has a surface temperature of about 6,000 K. At what wavelength does the Sun's spectrum peak? Answer:

$$\lambda_{\max} = 2.9 \times 10^{-3}/6000 \sim 0.5\,\mu = 500\,\mathrm{nm} \tag{9.21}$$

[10] The worldwide nuclear energy generation is about 350 GW. Thus, an area of 5,000 m^2 on the Sun generates this amount

The temperature derived from the peak wavelength is called *Wien Temperature*, the temperature derived from the difference of intensity between two wavelengths (=color) is called *Color temperature*, etc. To measure color, a filter system must be defined. The most commonly used system is the UBV-system (Table 9.3) which has three bands that are located in the UV (U), blue (B), and visual (V) to measure the intensity I_V. The color of a star is measured by comparing its magnitude through one filter (e.g., red) with its magnitude through another (e.g., blue).

Table 9.3 Central wavelength and bandwidth of the UBVRI filter set

Name	Meaning	Centralλ	Bandwidth [nm]
U	Ultraviolet	360	66
B	Blue	440	98
V	Visual (green)	550	87
R	Red	700	207
I	Infrared	900	231

For example, m_V means the magnitude measured with the V filter. Therefore, instead of determining temperatures from the comparison of the spectrum of a star with the Planck law, one can use, e.g., color indices. The value $B-V$ will be (see, e.g., Table 9.4):

- Positive for the cooler star, because it is brighter in V than in B (blue). If the cool star is brighter in V, it means that its magnitude has a lower value and, therefore, B-V is positive.
- Negative for the hotter star. The hotter star is brighter in B than in V, $m_B < m_V$ and B-V$<$0.

Table 9.4 B-V colors and effective temperatures of some stars

Star	B-V	Effective T
Sun	$+0.6$	5,800 K
Vega	0.0	10,000 K
Spica	-0.2	23,000 K
Antares	$+1.8$	3,400K

Habitable planetary systems need host stars in the spectral range F, G, or K. By measuring the colors of stars, these spectral types can be determined easily. F and G stars are yellow; K stars are orange.

Bibliography

1. G. Abramson. Ecological model of extinctions. *Physical Review E*, 55:785–788, January 1997.
2. L. W. Alvarez, W. Alvarez, F. Asaro, and H. V. Michel. Extraterrestrial cause for the cretaceous tertiary extinction. *Science*, 208:1095, 1980.
3. S. A. Amundson and D. J. Chen. Ionizing radiation-induced mutation of human cells with different DNA repair capacities. *Advances in Space Research*, 18:119–126, 1996.
4. P. D. Archer, A. A. Pavlov, and P. H. Smith. Atmospheric methane and martian climate. *LPI Contributions*, 1353:3193, July 2007.
5. S. Arrhenius.Panpermy: The tranmission of life from Star to Star. *Scientific American*, 196:196,1907.
6. D. J. Asher, M. Bailey, V. Emel'Yanenko, and W. Napier. Earth in the cosmic shooting gallery. *The Observatory*, 125:319–322, October 2005.
7. Z. Balog, G. H. Rieke, K. Y. L. Su, J. Muzerolle, and E. T. Young. Spitzer MIPS 24 μm Detection of Photoevaporating Protoplanetary Disks. *Astrophysical Journal Letters*, 650:L83–L86, October 2006.
8. M. Banit, M. A. Ruderman, J. Shaham, and J. H. Applegate. Formation of Planets around Pulsars. *Astrophysical Journal*, 415:779, October 1993.
9. A. Bar-Nun and V. Dimitrov. Methane on Mars: A product of H_2O photolysis in the presence of CO. *Icarus*, 188:543–545, June 2007.
10. E. S. Barghoorn and J. W. Schopf. Microorganisms three billion years old from the precambrian of South Africa. *Science*, 152:758–763, May 1966.
11. A. R. Basu, M. I. Petaev, R. J. Poreda, S. B. Jacobsen, and L. Becker. Chondritic meteorite fragments associated with the permian-triassic boundary in Antarctica. *Science*, 302:1388–1392, November 2003.
12. L. Becker, A. Shukolyukov, C. Macaissic, G. Lugmair, and R. Poreda. ET Extraterrestrial Chromium at the Graphite Peak P/Tr Boundary and in the Bedout Impact Melt Breccia. In S. Mackwell and E. Stansbery, editors, *37th Annual Lunar and Planetary Science Conference*, volume 37 of *Lunar and Planetary Institute Conference Abstracts*, page 2321, March 2006.
13. G. F. Benedict, B. E. McArthur, G. Gatewood, E. Nelan, W. D. Cochran, A. Hatzes, M. Endl, R. Wittenmyer, S. L. Baliunas, G. A. H. Walker, S. Yang, M. Kürster, S. Els, and D. B. Paulson. The Extrasolar Planet ε Eridani b: Orbit and Mass. *The Astronomical Journal*, 132:2206–2218, November 2006.
14. M. P. Bernstein. Extraterrestrial organic chemistry: From the interstellar medium to the origins of life. *Advances in Space Research*, 30:1393–1393, 2002.
15. C. Bertout, G. Basri, and S. Cabrit. The Classical T Tauri stars - Future Solar Systems? In C. P. Sonett, M. S. Giampapa, and M. S. Matthews, editors, *The Sun in Time*, pages 682–709, 1991.
16. V. Bhalerao and M. N. Vahia. Mass limit on Nemesis. *Bulletin of the Astronomical Society of India*, 33:27–+, March 2005.

17. R. Bielski and M. Tencer. A possible path to the RNA world: Enantioselective and diastereoselective purification of ribose. *Origins of Life and Evolution of the Biosphere*, 37:167–175, April 2007.
18. K. Biemann, J. Oro, P. Toulmin, III, L. E. Orgel, A. O. Nier, D. M. Anderson, D. Flory, A. V. Diaz, D. R. Rushneck, and P. G. Simmonds. The search for organic substances and inorganic volatile compounds in the surface of Mars. *Journal of Geophys. Research*, 82:4641–4658, September 1977.
19. R. P. Binzel. The torino impact hazard scale. *Planetary and Space Science*, 48:297–303, April 2000.
20. J. Birn. On the stability of the planetary system. *Astronomy and Astrophysics*, 24:283, April 1973.
21. D. I. Black. Cosmic ray effects and faunal extinctions at geomagnetic field reversals. *Earth and Planetary Science Letters*, 3:225–236, 1967.
22. D. L. Block, F. Bournaud, F. Combes, R. Groess, P. Barmby, M. L. N. Ashby, G. G. Fazio, M. A. Pahre, and S. P. Willner. An almost head-on collision as the origin of two off-centre rings in the Andromeda galaxy. *Nature*, 443:832–834, October 2006.
23. W. F. Bottke and A. Morbidelli. The asteroid and comet impact flux in the terrestrial planet region: A brief history of the last 4.6 Gy. *LPI Contributions*, 1320:18–19, May 2006.
24. G. C. Brainhard. The Influence of various irradiances of artificial light, twilight, and moonlight on the suppression of pineal melatonin content in the Syrian Hamster. *Pineal Research*, 1984.
25. R. T. Brown. Ionospheric Effects of Cosmic γ-ray Bursts. *Nature*, 246:83–84, November 1973.
26. A. P. Buccino, G. A. Lemarchand, and P. J. D. Mauas. Ultraviolet radiation constraints around the circumstellar habitable zones. *Icarus*, 183:491–503, August 2006.
27. D. Buhl. Chemical costituents of interstellar clouds. *Nature* 234:332,1971.
28. M. J. Burchell. W(h)ither the Drake equation? *International Journal of Astrobiology*, 5:243–250, December 2006.
29. A. J. Burgasser. The physical properties of HD 3651B: An extrasolar nemesis? *Astrophysical Journal*, 658:617–621, March 2007.
30. A. J. Burgasser, J. D. Kirkpatrick, M. R. McGovern, I. S. McLean, L. Prato, and I. N. Reid. S Orionis 70: Just a foreground field brown dwarf? *Astrophysical Journal*, 604:827–831, April 2004.
31. A. G.Cairns-Smith. Genetic takeover and the mineral origins of life. New York, Cambridge University Press,1982.
32. A. C. Cameron, K. Horne, A. Penny, and D. James. Probable detection of starlight reflected from the giant planet orbiting τ Boötis. *Nature*, 402:751–755, December 1999.
33. R. M. Canup. Simulations of a late lunar-forming impact. *Icarus*, 168:433–456, April 2004.
34. R. W. Carlson, P. Drossart, T. Encrenaz, P. R. Weissman, J. Hui, and M. Segura. Temperature, size, and energy of the Shoemaker-Levy 9 G-impact fireball. *Icarus*, 128:251–274, August 1997.
35. R. Cavicchioli. Extremophiles and the search for extraterrestrial life. *Astrobiology*, 2:281–292, August 2002.
36. C. R. Chapman and D. Morrison. Impacts on the Earth by asteroids and comets: Assessing the hazard. *Nature*, 367:33–40, January 1994.
37. C. R. Chapman and B. M. Mulligan. NEO Impact Hazard Scales in the Context of Other Hazard Scales. In *Bulletin of the American Astronomical Society*, page 861, September 2002.
38. S. R. Chesley, P. W. Chodas, A. Milani, G. B. Valsecchi, and D. K. Yeomans. Quantifying the risk posed by potential Earth impacts. *Icarus*, 159:423–432, October 2002.
39. M. M. Ćirković. The temporal aspect of the drake equation and SETI. *Astrobiology*, 4:225–231, June 2004.

40. J. N. Clarke. Extraterrestrial intelligence and galactic nuclear activity. *Icarus*, 46:94–96, April 1981.
41. C. S. Cockell. Life on Venus. *Planetary and Space Science*, 47:1487–1501, December 1999.
42. C. S. Cockell, P. Lee, P. Broady, D. S. S. Lim, G. R. Osinski, J. Parnell, C. Koeberl, L. Pesonen, and J. Salminen. Effects of asteroid and comet impacts on habitats for lithophytic organisms - A synthesis. *Meteoritics and Planetary Science*, 40:1901–1914, December 2005.
43. J. F. Cooper. Habitability in high radiation environments: The case for gaia at Europa. *AGU Fall Meeting Abstracts*, pages A1125+, December 2004.
44. A. C. M. Correia and J. Laskar. The four final rotation states of Venus. *Nature*, 411:767–770, June 2001.
45. A. C. M. Correia and H. Silva. Capture probability of NEOs by the Earth-Moon system. In *IAU Symposium*, August 2006.
46. T. J. Cox and A. Loeb. The collision between the Milky Way and Andromeda. *ArXiv e-prints*, 705, May 2007.
47. J. Cronin and J. Reisse. *Chirality and the Origin of Homochirality*, page 473. Lectures in Astrobiology, Volume 1, 2005.
48. W. Cunningham, M.A. Cunningham, and B.W.Saigo, *Environmental Science*. McGraw Hill, 2005.
49. M. R. Cunningham, P. A. Jones, P. D. Godfrey, D. M. Cragg, I. Bains, M. G. Burton, P. Calisse, N. H. M. Crighton, S. J. Curran, T. M. Davis, J. T. Dempsey, B. Fulton, M. G. Hidas, T. Hill, L. Kedziora-Chudczer, V. Minier, M. B. Pracy, C. Purcell, J. Shobbrook, and T. Travouillon. A search for propylene oxide and glycine in Sagittarius B2 (LMH) and Orion. *Monthly Notices*, 376:1201–1210, April 2007.
50. M. W. DeLaubenfels. Dinosaur extinctions: One more hypothesis. *jounal of Palentology*, 30(1):207-218, January 1956.
51. C. S. M. Doake. Climatic change and geomagnetic field reversals: A statistical correlation. *Earth and Planetary Science Letters*, 38:313–318, February 1978.
52. A. Dressler and D. O. Richstone. Stellar dynamics in the nuclei of M31 and M32 - Evidence for massive black holes? *Astrophysical Journal*, 324:701–713, January 1988.
53. J. Dubinski. The great Milky Way -Andromeda collision. *Sky and Telescope*, 112(4):30–36, 2006.
54. M. Eigen and P. Schuster. *The Hypercycle-A Pricipal of Natural Self- Organization*, Springer-Verlag,Berlin, 1979.
55. J. Ellis and D. N. Schramm. Could a nearby supernova explosion have caused a Mass extinction? *Proceedings of the National Academy of Science*, 92:235–238, January 1995.
56. K. Fadel, E. V. Vashenyuk, and A. S. Kirillov. Ozone depletion in the middle atmosphere during solar proton events in October 2003. *Advances in Space Research*, 38:1881–1886, 2006.
57. P. Farinella, L. Foschini, C. Froeschlé, R. Gonczi, T. J. Jopek, G. Longo, and P. sMichel. Probable asteroidal origin of the tunguska cosmic body. *Astronomy and Astrophysics*, 377:1081–1097, October 2001.
58. L. K. Fenton, P. E. Geissler, and R. M. Haberle. Global warming and climate forcing by recent albedo changes on Mars. *Nature*, 446:646–649, April 2007.
59. J. P. Ferris, A. R. Hill, R. Liu, and L. E. Orgel. Synthesis of long prebiotic oligomers on mineral surfaces. *Nature*, 381:59–61, May 1996.
60. D. A. Fischer, G. W. Marcy, R. P. Butler, S. S. Vogt, G. Laughlin, G. W. Henry, D. Abouav, K. M. G. Peek, J. T. Wright, J. A. Johnson, C. McCarthy, and H. Isaacson. Five planets orbiting 55 Cancri. *Astrophysical Journal*, 675:790–801, March 2008.
61. M. P. Fitzgerald. The distribution of interstellar reddening material. *Astronology. Journal*, 73:983, December 1968.
62. E. B. Ford, S. Seager, and E. L. Turner. On the red edge an optical biomarker for detecting extrateresstrial plants. *AGU Fall Meeting Abstracts*, pages A6+, December 2005.
63. F. Forget and R. T. Pierrehumbert. Warming early Mars with carbon dioxide clouds that scatter infrared radiation. *Science*, 278:1273–+, November 1997.

64. S. Franck, W. von Bloh, C. Bounama, M. Steffen, D. Schönberner, and H. J. Schellnhuber. Limits of photosynthesis in extrasolar planetary systems for earth-like planets. *Advances in Space Research*, 28:695–700, 2001.
65. P. C. Frisch. The local bubble and interstellar material near the Sun. *Space Science Reviews*, page 124, July 2007.
66. B. Fuchs, D. Breitschwerdt, M. A. de Avillez, C. Dettbarn, and C. Flynn. The search for the origin of the Local Bubble redivivus. *Monthly Notices*, 373:993–1003, December 2006.
67. Y. Funato, J. Makino, P. Hut, E. Kokubo, and D. Kinoshita. The formation of Kuiper-belt binaries through exchange reactions. *Nature*, 427:518–520, February 2004.
68. H. E. Garcia, T. P. Boyer, S. Levitus, R. A. Locarnini, and J. Antonov. On the variability of dissolved oxygen and apparent oxygen utilization content for the upper world ocean: 1955 to 1998. *Geophysical Research Letter*, 32:9604–+, May 2005.
69. N. Gehrels, C. H. Jackman, J. K. Cannizzo, B. J. Mattson, and C. M. Laird. Ozone Depletion from Nearby Supernovae. In *Bulletin of the American Astronomical Society*, volume 34 of *Bulletin of the American Astronomical Society*, page 1267, December 2002.
70. M. S. Giampapa and C. L. Imhoff. The Ambient Radiation Field of Young Solar Systems Ultraviolet and X-ray Emission from T Tauri Stars. In D. C. Black and M. S. Matthews, editors, *Protostars and Planets II*, pages 386–404, 1985.
71. L. Gizon and A. C. Birch. Local Helioseismology. *Living Reviews in Solar Physics*, 2:6, November 2005.
72. B. Gladman, D. Dane Quinn, P. Nicholson, and R. Rand. Synchronous Locking of Tidally Evolving Satellites. *Icarus*, 122:166–192, July 1996.
73. J. Glanz. Hints of a planet orbiting Sunlike Star - 51-PEGASI. *Science*, 270:375, October 1995.
74. G. A. Glatzmaier and P. H. Roberts. A three-dimensional self-consistent computer simulation of a geomagnetic field reversal. *Nature*, 377:203, September 1995.
75. A. Glikson. Asteroid/comet impact clusters, flood basalts and mass extinctions: Significance of isotopic age overlaps. *Earth and Planetary Science Letters*, 236:933–937, August 2005.
76. D. Goldsmith. *The Hunt for Life on Mars*, Penguin Books, New york,1997.
77. G. Gonzalez. Spectroscopic Analyses of the planetary system candidates: 55 Cnc, 51 Peg, 47 UMa, 70 Vir, and HD 114762. *ArXiv Astrophysics e-prints*, September 1996.
78. D. O. Gough. Solar interior structure and luminosity variations. *Solar Physics*, 74:21–34, November 1981.
79. K. Goździewski, M. Konacki, and A. Wolszczan. Long-term stability and dynamical environment of the PSR 1257+12 planetary system. *Astrophysical Journal*, 619:1084–1097, February 2005.
80. D. F. Gray and S. L. Baliunas. Magnetic activity variations of epsilon Eridani. *Astrophysical Journal*, 441:436–442, March 1995.
81. J. S. Greaves, M. C. Wyatt, W. S. Holland, and W. R. F. Dent. The debris disc around τ Ceti: a massive analogue to the Kuiper Belt. *Monthly Notices*, 351:L54–L58, July 2004.
82. I. A. Grenier, J.-M. Casandjian, and R. Terrier. Unveiling Extensive Clouds of Dark Gas in the Solar Neighborhood. *Science*, 307:1292–1295, February 2005.
83. J. Gribbin. Martian climate: Past, present and future. *Astronomy*, 5:18–24, 1977.
84. J.-M. Grießmeier, A. Stadelmann, T. Penz, H. Lammer, F. Selsis, I. Ribas, E. F. Guinan, U. Motschmann, H. K. Biernat, and W. W. Weiss. The effect of tidal locking on the magnetospheric and atmospheric evolution of "Hot Jupiters". *Astronomy and Astrophysics*, 425:753–762, October 2004.
85. D. H. Grinspoon and M. A. Bullock. Impact Induced Climate Change on Venus: The Role of Large Comets. In *Bulletin of the American Astronomical Society*, page 1119, October 2000.
86. H. Gursky and D. A. Schwartz. Extragalactic X-ray sources. *Astronomy and Astrophysics*, 15:541–568, 1977.
87. J. B. Haldane. The origins of life. *New Biology*, 1954.
88. I. Halevy, M. T. Zuber, and D. P. Schrag. A Sulfur Dioxide Climate Feedback on Early Mars. *Science*, 318:1903, December 2007.

89. A. Hamdouni, A. Barbe, P. Demoulin, and R. Zander. Retrieval of ozone vertical column amounts from ground-based high resolution infrared solar spectra. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 57:11–22, January 1997.
90. B. M. S. Hansen. Stellar Collisions and Pulsar Planets. In M. M. Shara, editor, *Stellar Collisions, Mergers and their Consequences*, volume 263 of *Astronomical Society of the Pacific Conference Series*, page 221, 2002.
91. A. Hanslmeier. *The Sun and Space Weather.* Springer. Astrophysics and Space Science Library, vol. 237, 2007.
92. A. W. Harris and W. R. Ward. Dynamical constraints on the formation and evolution of planetary bodies. *Annual Review of Earth and Planetary Sciences*, 10:61–108, 1982.
93. D. E. Harris, J. A. Biretta, W. Junor, E. S. Perlman, W. B. Sparks, and A. S. Wilson. Flaring X-Ray Emission from HST-1, a Knot in the M87 Jet. *Astrophysical Journal Letters*, 586:L41–L44, March 2003.
94. N. W. Harris and M. E. Bailey. Dynamical evolution of cometary asteroids. *Monthly Notices*, 297:1227–1236, July 1998.
95. M. H. Hart. Habitable Zones about Main Sequence Stars. *Icarus*, 37:351–357, January 1979.
96. W. K. Hartmann and D. R. Davis. Satellite-sized planetesimals and lunar origin. *Icarus*, 24:504–514, April 1975.
97. J. D. Hays, J. Imbrie, and N. J. Shackleton. Variations in the Earth's orbit: pacemaker of the ice ages. *Science*, 194:1121–1132, December 1976.
98. R. M. Hazen and D. W. Deamer. Hydrothermal reactions of pyruvic acid: Synthesis, selection, and self-assembly of amphiphilic molecules. *Origins of Life and Evolution of the Biosphere*, 37:143–152, April 2007.
99. M. Heath. The Forest-Habitability of Earthlike Planets. In L. R. Doyle, editor, *Circumstellar Habitable Zones*, page 445, 1996.
100. M. J. Heath and L. R. Doyle. From Near-Synchronously Rotating Planets to Tidal Lock: A New Class of Habitable Planets Examined for Forest Habitability. In R. Norris and F. Stootman, editors, *Bioastronomy 2002: Life Among the Stars*, volume 213 of *IAU Symposium*, page 225, June 2004.
101. W. M. Heckl.*Astrobiology, The Quest for the Conditions of Life*,Springer, 2000
102. A. R. Hildebrand and W. V. Boynton. On the Location of the K/T Boundary Impact Site. In *Lunar and Planetary Institute Conference Abstracts*, volume 21 of *Lunar and Planetary Institute Conference Abstracts*, page 512, March 1990.
103. J. G. Hills. Dynamical constraints on the mass and perihelion distance of Nemesis and the stability of its orbit. *Nature*, 311:636–638, October 1984.
104. J. G. Hills, C. L. Mader, M. P. Goda, and M. S. Warren. Down-to-Earth Astronomy: Tsunami from Asteroid-Comet Impacts. In *Bulletin of the American Astronomical Society*, page 1260, December 1997.
105. C. Hofmann, V. Courtillot, G. Féraud, P. Rochette, G. Yirgu, E. Ketefo, and R. Pik. Timing of the Ethiopian flood basalt event and implications for plume birth and global change. *Nature*, 389:838, October 1997.
106. C. Hofmann, G. Féraud, and V. Courtillot. $^{40}Ar/^{39}Ar$ dating of mineral separates and whole rocks from the Western Ghats lava pile: further constraints on duration and age of the Deccan traps. *Earth and Planetary Science Letters*, 180:13–27, July 2000.
107. F. Hoyle. *The Black Cloud*, Signet, New York, 1959.
108. F. Hoyle and N. C. Wickramasinghe. Origin and nature of carbonaceous material in the galaxy. *Nature*, 270:701–703, December 1977.
109. J. D. Huba, H. P. Warren, G. Joyce, X. Pi, B. Iijima, and C. Coker. Global response of the low-latitude to midlatitude ionosphere due to the Bastille Day flare. *Geophysical Research Letter*, 32:15103, August 2005.
110. E. Hubble. No. 376. A spiral nebula as a stellar system. Messier 31. *Contributions from the Mount Wilson Observatory / Carnegie Institution of Washington*, 376:1–55, 1929.

111. C. Huber and G. Waechtershaeuser. Peptides by activation of amino acids with CO on (Ni,Fe)S surfaces: Implications for the origin of life. *Science*, 281:670, July 1998.
112. S. Ida, R. M. Canup, and G. R. Stewart. Lunar accretion from an impact-generated disk. *Nature*, 389:353–357, September 1997.
113. R. Ignace and M. Giroux. Photoevaporation of Planets around Massive Stars. In R. Ignace and K. G. Gayley, editors, *The Nature and Evolution of Disks Around Hot Stars*, volume 337 of *Astronomical Society of the Pacific Conference Series*, page 241, November 2005.
114. H. Imanaka, B. N. Khare, and C. P. McKay. UV shielding of early Earth by N2/CH4/CO2 organic haze. In *35th COSPAR Scientific Assembly*, volume 35 of *COSPAR, Plenary Meeting*, page 1076, 2004.
115. T. Ito and K. Tanikawa. Long-term integrations and stability of planetary orbits in our Solar system. *Monthly Notices*, 336:483–500, October 2002.
116. J. F. Kasting. Runaway and moist greenhouse atmospheres and the evolution of earth and Venus. *Icarus*, 74:472–494, June 1988.
117. J. F. Kasting, D. P. Whitmire, and R. T. Reynolds. Habitable zones around main sequence stars. *Icarus*, 101:108–128, January 1993.
118. T. Kato and A. E. Ringwood. Was the Moon Formed from the Mantle of a Martian-sized Planetesimal? In *Lunar and Planetary Institute Conference Abstracts*, volume 20 of *Lunar and Planetary Institute Conference Abstracts*, page 510, March 1989.
119. G. Keller. Biotic effects of impacts and volcanism. *Earth and Planetary Science Letters*, 215:249–264, October 2003.
120. G. Keller. Biotic effects of late Maastrichtian mantle plume volcanism: implications for impacts and mass extinctions. *Lithos*, 79:317–341, February 2005.
121. B. N. Khare, C. Sagan, E. L. Bandurski, and B. Nagy. Ultraviolet-photoproduced organic solids synthesized under simulated Jovian conditions: Molecular analysis. *Science*, 199:1199–1201, March 1978.
122. N. Y. Kiang, A. Segura, G. Tinetti, Govindjee, R. E. Blankenship, M. Cohen, J. Siefert, D. Crisp, and V. S. Meadows. Spectral signatures of photosynthesis. II. coevolution with other stars and the atmosphere on extrasolar worlds. *Astrobiology*, 7:252–274, February 2007.
123. J. W. Kirchner. The Gaia hypothesis: Can it be tested? *Reviews of Geophysics*, 27:223–235, 1989.
124. J. D. Kirkpatrick. A New Census of the Solar Neighborhood. In C. E. Woodward, M. D. Bicay, and J. M. Shull, editors, *Tetons 4: Galactic Structure, Stars and the Interstellar Medium*, volume 231 of *Astronomical Society of the Pacific Conference Series*, page 17, 2001.
125. A. A. Kisselev, Y. N. Gnedin, E. A. Grosheva, N. A. Shakht, D. L. Gorshanov, and M. Y. Piotrovich. The supermassive black hole at the center of our galaxy: Determination of its main physical parameters. *Astronomy Reports*, 51:100–108, February 2007.
126. R. W. Klebesadel, I. B. Strong, and R. A. Olson. Observations of Gamma-Ray Bursts of Cosmic Origin. *Astrophysical Journal Letters*, 182:L85+, June 1973.
127. N. S. Kovalenko, K. I. Churyumov, and Y. G. Babenko. Centaur Population As A Source Of NEO's And Active Comets – Dynamical Evolution Modeling. In *IAU Symposium*, August 2006.
128. A. T. Kovalev, I. V. Nemchinov, and V. V. Shuvalov. Ionospheric and magnetospheric disturbances caused by impacts of small comets and asteriods. *Solar System Research*, 40:57–67, January 2006.
129. L. Kresák. The Tunguska catastrophe and comet Encke. *Kozmos*, 9:163–166, 1978.
130. D.A. Kring. Air balst produced by the Meteor Crater impact event and a reconstruction of the affected environment. *Meteoritics and Planetary Journal*, 32:517, 1997.
131. Y.-J. Kuan, S. B. Charnley, H.-C. Huang, W.-L. Tseng, and Z. Kisiel. Interstellar Glycine. *Astrophysical Journal*, 593:848–867, August 2003.
132. O. L. Kuskov and V. A. Kronrod. Internal structure of Callisto: Evidence for subsurface ocean. In *36th COSPAR Scientific Assembly*, volume 36 of *COSPAR, Plenary Meeting*, page 761, 2006.

133. P.-O. Lagage, C. Doucet, E. Pantin, E. Habart, G. Duchêne, F. Ménard, C. Pinte, S. Charnoz, and J.-W. Pel. Anatomy of a flaring proto-planetary disk around a young intermediate-mass star. *Science*, 314:621–623, October 2006.
134. S. A. Lamb, P. F. Miles, N. C. Hearn, and D. R. Schwenk. Simulations of the Evolution of the Gas, and the Stellar and Dark Matter Distributions, Caused by the Collision of the Galaxy Pair NGC 3395/6 (Arp 270). In *Bulletin of the American Astronomical Society*, volume 38 of *Bulletin of the American Astronomical Society*, page 373, September 2006.
135. W. D. Langer and A. E. Glassgold. Silicon chemistry in interstellar clouds. *Astrophysical Journal*, 352:123–131, March 1990.
136. J. Laskar. A numerical experiment on the chaotic behaviour of the solar system. *Nature*, 338:237, March 1989.
137. J. Laskar. Large-scale chaos in the solar system. *Astronomy and Astrophysics*, 287:L9–L12, July 1994.
138. J. Laskar, F. Joutel, and P. Robutel. Stabilization of the earth's obliquity by the moon. *Nature*, 361:615–617, February 1993.
139. T. J. W. Lazio, J. Fischer, and J. M. Cordes. Searches for Pulsar Planetary Systems. In R. Norris and F. Stootman, editors, *Bioastronomy 2002: Life Among the Stars*, volume 213 of *IAU Symposium*, page 101, June 2004.
140. J,-M. Lehn, *Supramolecular Chemistry, Concepts and Perspectives*, Weley VCH, 1995.
141. H. Levison and M. J. Duncan. Transport from TNOs to NEOs. In *IAU Symposium*, August 2006.
142. H. F. Levison, M. J. Duncan, K. Zahnle, M. Holman, and L. Dones. NOTE: Planetary Impact Rates from Ecliptic Comets. *Icarus*, 143:415–420, February 2000.
143. J. S. Lewis. *Comet and Asteriod Impact Hazards on a Populated Earth : Computer Modeling*. San Diego, Academic Press, 2000.
144. D. N. C. Lin and J. Papaloizou. On the tidal interaction between protoplanets and the proto-planetary disk. III – Orbital migration of protoplanets. *Astrophysical Journal*, 309:846–857, October 1986.
145. C. H. Lineweaver, Y. Fenner, and B. K. Gibson. The galactic habitable zone and the age distribution of complex life in the Milky Way. *Science*, 303:59–62, January 2004.
146. L. E. Lisiecki and M. E. Raymo. A Pliocene-Pleistocene stack of 57 globally distributed benthic δ^{18}O records. *Paleoceanography*, 20:1003, January 2005.
147. A. Loeb. Observable signatures of a black hole ejected by gravitational-radiation recoil in a Galaxy merger. *Physical Review Letters*, 99(4):041103, July 2007.
148. M. López-Puertas, B. Funke, S. Gil-López, T. von Clarmann, G. P. Stiller, M. Höpfner, S. Kellmann, H. Fischer, and C. H. Jackman. Observation of NO enhancement and ozone depletion in the Northern and Southern Hemispheres after the October–November 2003 solar proton events. *Journal of Geophysical Research (Space Physics)*, 110(9):9, September 2005.
149. J. E. Lovelock and L. Margulis. Atmospheric homeostasis by and for the biosphere: The gaia hypothesis. *Tellus*, 26:2, August 1974.
150. K. L. Luhman, L. Adame, P. D'Alessio, N. Calvet, L. Hartmann, S. T. Megeath, and G. G. Fazio. Discovery of a planetary-mass brown dwarf with a circumstellar disk. *Astrophysical Journal Letters*, 635:L93–L96, December 2005.
151. J. R. Lyons, T. J. Ahrens, and A. R. Vasavada. Global Effects of Impact Ejecta from the K-T Bolide. In *Lunar and Planetary Institute Conference Abstracts*, volume 28 of *Lunar and Planetary Institute Conference Abstracts*, page 857, March 1997.
152. R. L. Mancinelli. Peroxides and the survivability of microorganisms on the surface of Mars. *Advances in Space Research*, 9:191–195, 1989.
153. M. Marengo, S. T. Megeath, G. G. Fazio, K. R. Stapelfeldt, M. W. Werner, and D. E. Backman. A Spitzer IRAC search for substellar companions of the debris disk Star ε Eridani. *Astrophysical Journal*, 647:1437–1451, August 2006.
154. L. Margulis and J. E. Lovelock. Biological modulation of the Earth's atmosphere. *Icarus*, 21:471, April 1974.
155. L. Margulis and J. E. Lovelock. Is Mars a spaceship, too. *Natural History*, 85:86–90, July 1976.

156. W.P. Marshall and, C.R. Marshall. Sudden and gradual molluscan extinctions in the latest Cretaceous of Western European Tethys. *AGU Fall Meeting Abstracts*, 1996.
157. F. Marzari and S. J. Weidenschilling. Eccentric extrasolar planets: The jumping Jupiter model. *Icarus*, 156:570–579, April 2002.
158. M. Massi. Stellar flaring periodicities. *Memorie della Societa Astronomica Italiana*, 78:247, 2007.
159. I. Matsuyama, D. Johnstone, and N. Murray. Halting planet migration by photoevaporation from the central source. *Astrophysical Journal Letters*, 585:L143–L146, March 2003.
160. L. Mayer, S. Kazantzidis, P. Madau, M. Colpi, T. Quinn, and J. Wadsley. Rapid formation of supermassive black hole binaries in Galaxy mergers with gas. *Science*, 316:1874, June 2007.
161. M. Mayor, D. Queloz, G. Marcy, P. Butler, R. Noyes, S. Korzennik, M. Krockenberger, P. Nisenson, T. Brown, T. Kennelly, C. Rowland, S. Horner, G. Burki, M. Burnet, and M. Kunzli. 51 Pegasi. *IAU Circ.*, 6251:1–+, October 1995.
162. C. P. McKay, R. D. Lorenz, and J. I. Lunine. Analytic solutions for the antigreenhouse effect: Titan and the Early Earth. *Icarus*, 137:56–61, January 1999.
163. A. L. Melott, B. S. Lieberman, C. M. Laird, L. D. Martin, M. V. Medvedev, B. C. Thomas, J. K. Cannizzo, N. Gehrels, and C. H. Jackman. Did a gamma-ray burst initiate the late Ordovician mass extinction? *International Journal of Astrobiology*, 3:55–61, January 2004.
164. M. Menichella, P. Paolicchi, and P. Farinella. The main belt as a source of near-Earth asteroids. *Earth Moon and Planets*, 72:133–149, 1996.
165. K. M. Meyer, L. R. Kump, and A. Ridgwell. Global warmth and nutrient trapping enhance end-paleozoic euxinia in an Earth system model. *AGU Fall Meeting Abstracts*, pages C4+, December 2007.
166. S. L. Miller. Production of amino acids under possible primitive Earth conditions. *Science*, 117:528, 1953.
167. S. L. Miller and H.C.Urey. Organicr compound synthesis on the primitive Earth *Science*,130:245,1959.
168. M. A. Mischna, J. F. Kasting, A. Pavlov, and R. Freedman. Influence of carbon dioxide clouds on early martian climate. *Icarus*, 145:546–554, June 2000.
169. R. Moreno, A. Marten, and H. E. Matthews. Jupiter's Stratosphere 10 Years after the Collision of Comet Shoemaker-Levy 9. In *Bulletin of the American Astronomical Society*, volume 37 of *Bulletin of the American Astronomical Society*, page 660, August 2005.
170. A. Morlok, M. Köhler, M. Anand, C. Kirk, and M. M. Grady. Dust from collisions in circumstellar disks: Similarities to meteoritic materials? *Meteoritics Planetary Science, Vol. 41, Supplement, Proceedings of 69th Annual Meeting of the Meteoritical Society, held August 6–11, 2006 in Zurich, Switzerland*, 41:5135, September 2006.
171. D. Morrison, editor. *The Spaceguard Survey*, 1992.
172. F. Mullally and D. Winget. A Possible Planet Around a White Dwarf. In *Bulletin of the American Astronomical Society*, volume 38 of *Bulletin of the American Astronomical Society*, pages 1129–+, December 2006.
173. R. Muller. The Solar Granulation. In A. Hanslmeier and M. Messerotti, editors, *Motions in the Solar Atmosphere*, volume 239 of *Astrophysics and Space Science Library*, pages 35–70, 1999.
174. R. Muller. *Ice ages and astronomical causes: data, spectral analysis, and mechanisms*. Richard A. Muller and Gordon J. MacDonald, London, New York, Springer, c2000. Springer-Praxis books in environmental sciences, 2000.
175. N. Murray, B. Hansen, M. Holman, and S. Tremaine. Migrating Planets. *Science*, 279:69–+, January 1998.
176. W. M. Napier. Evidence for cometary bombardment episodes. *Monthly Notices*, 366:977–982, March 2006.
177. R. P. Nelson, J. C. B. Papaloizou, F. Masset, and W. Kley. The migration and growth of protoplanets in protostellar discs. *Monthly Notices*, 318:18–36, October 2000.

178. I. V. Nemchinov, M. P. Popova, L. P. Shubadeeva, V. V. Shuvalov, and V. V. Svetsov. Effects of Hydrodynamics and Thermal Radiation in the Atmosphere after Comet Impacts. In *Lunar and Planetary Institute Conference Abstracts*, pages 1067–1068, March 1993.
179. E. W. Nester, D.G. Anderson, C.E. Roberts, N.N. Pearshall, and M.T. Nester, *Microbiology*, McGraw Hill, 2001.
180. M. E. J. Newman. A Model of Mass Extinction. In *eprint arXiv:adap-org/9702003*, page 2003, February 1997.
181. M. Noble, Z. E. Musielak, and M. Cuntz. Orbital stability of terrestrial planets inside the habitable zones of extrasolar planetary systems. *Astrophysical Journal*, 572:1024–1030, June 2002.
182. D. Z. Oehler, F. Robert, A. Meibom, S. Mostefaoui, M. Selo, M. Walter, K. Sugitani, A. Allwood, K. Mimura, and E. K. Gibson. "Nano" Scale Biosignatures and the Search for Extraterrestrial Life. In *Lunar and Planetary Institute Conference Abstracts, volume 39 of Lunar and Planetary Institute Conference Abstracts*, pages 1303–+, March 2008.
183. L. Ong, E. Asphaug, and C. Plesko. Water Delivered to the Moon by Comet Impacts. In S. Mackwell and E. Stansbery, editors, *37th Annual Lunar and Planetary Science Conference*, page 2450, March 2006.
184. T. C. Onstott, D. McGown, J. Kessler, B. S. Lollar, K. K. Lehmann, and S. M. Clifford. Martian CH_4: Sources, flux, and detection. *Astrobiology*, 6:377–395, April 2006.
185. T. Owen and A. Bar-Nun. Comets, impacts and atmospheres. *Icarus*, 116:215–226, 1995.
186. E. N. Parker. The occasional reversal of the geomagnetic field. *Astrophysical Journal Letters*, 158:815, November 1969.
187. I. de Pater and J. Lissauer. *Planetary Sciences*, Cambridge University Press, 2001.
188. E. S. Perlman, D. E. Harris, J. A. Biretta, W. B. Sparks, and F. D. Macchetto. Month-timescale optical variability in the M87 Jet. *Astrophysical Journal Letters*, 599:L65–L68, December 2003.
189. C. B. Phillips and C. F. Chyba. Impact Gardening Rates on Europa: Comparison with Sputtering. In *Lunar and Planetary Institute Conference Abstracts, volume 32 of Lunar and Planetary Inst. Technical Report*, page 2111, March 2001.
190. E. Pierazzo and C. F. Chyba. Cometary delivery of biogenic elements to Europa. *Icarus*, 157:120–127, May 2002.
191. K. O. Pope, S. L. D'Hondt, and C. R. Marshall. Meteorite impact and the Mass extinction of species at the cretaceous/tertiary boundary. *Proceedings of the National Academy of Science*, 95:11028–11029, September 1998.
192. R. J. Poreda and L. Becker. Fullerenes and interplanetary dust at the permian-triassic boundary. *Astrobiology*, 3:75–90, January 2003.
193. D. Queloz, M. Mayor, L. Weber, A. Blécha, M. Burnet, B. Confino, D. Naef, F. Pepe, N. Santos, and S. Udry. The CORALIE survey for southern extra-solar planets. I. A planet orbiting the star Gliese 86. *Astrobiology*, 354:99–102, February 2000.
194. T. R. Quinn, S. Tremaine, and M. Duncan. A three million year integration of the earth's orbit Astrology Journal. , 101:2287–2305, June 1991.
195. E. Rabbow and G. Horneck. Mutagenic effects of solar UV-radiation on DNA. In P. Ehrenfreund, O. Angerer, and B. Battrick, editors, *ESA SP-496: Exo-/Astro-Biology*, pages 393–396, August 2001.
196. D. L. Rabinowitz, E. Bowell, E. M. Shoemaker, and K. Muinonen. The Population of Earth-Crossing Asteroids. In T. Gehrels, M. S. Matthews, and A. M. Schumann, editors, *Hazards Due to Comets and Asteroids*, page 285, 1994.
197. D. Raghavan, T. J. Henry, J. Subasavage, and T. Beaulieu. Two Suns in the Sky: Many Extrasolar Planets Orbiting Stars in Multiple Systems. In *Bulletin of the American Astronomical Society*, volume 35 of *Bulletin of the American Astronomical Society*, page 1271, December 2003.
198. M. R. Rampino. The galactic theory of Mass extinctions: An update. *Celestial Mechanics and Dynamical Astronomy*, 69:49–58, September 1997.

199. M. R. Rampino and R. B. Stothers. Flood basalt volcanism during the past 250 million years. *Science*, 241:663–668, August 1988.
200. M. R. Rampino and R. B. Stothers. Mass extinctions, comet impacts, and the Galaxy. *Highlights of Astronomy*, 11:246, 1998.
201. O. M. Raspopov, V. A. Dergachev, and T. Kolström. Hale cyclicity of Solar activity and its relation to climate variability. *Solar Physics*, 224:455–463, October 2004.
202. D. M. Raup. Magnetic reversals and mass extinctions. *Nature*, 314:341–343, March 1985.
203. D. M. Raup and J. J. Sepkoski, Jr. Periodicity of extinctions in the geologic past. *Proceedings of the National Academy of Science*, 81:801–805, 1984.
204. S. N. Raymond, A. M. Mandell, and S. Sigurdsson. Exotic Earths: Forming habitable worlds with giant planet migration. *Science*, 313:1413–1416, September 2006.
205. J. E. Reynolds. *Nature*, 48:477, 1893.
206. R. T. Reynolds, C. P. McKay, and J. F. Kasting. Europa, tidally heated oceans, and habitable zones around giant planets. *Advances in Space Research*, 7:125–132, 1987.
207. J. H. Rhee, I. Song, and B. Zuckerman. Warm dust in the terrestrial planet zone of a Sun-like Pleiades Star: Collisions between planetary embryos? *Astrophysical Journal*, 675:777–783, March 2008.
208. H. Rickman, J. A. Fernández, G. Tancredi, and J. Licandro. The cometary contribution to planetary impact rates. *Chemical Physics of Solid Surfaces*, pages 131–142, 2001.
209. J. H. Roberts and F. Nimmo. Stability of a Subsurface Ocean on Enceladus. In *Lunar and Planetary Institute Conference Abstracts, volume 38 of Lunar and Planetary Institute Conference Abstracts*, page 1429, March 2007.
210. R. D. Robinson, J. M. Wheatley, B. Y. Welsh, K. Forster, P. Morrissey, M. Seibert, R. M. Rich, S. Salim, T. A. Barlow, L. Bianchi, Y.-I. Byun, J. Donas, P. G. Friedman, T. M. Heckman, P. N. Jelinsky, Y.-W. Lee, B. F. Madore, R. F. Malina, D. C. Martin, B. Milliard, S. G. Neff, D. Schiminovich, O. H. W. Siegmund, T. Small, A. S. Szalay, and T. K. Wyder. GALEX observations of an energetic ultraviolet flare on the dM4e Star GJ 3685A. *Astrophysical Journal*, 633:447–451, November 2005.
211. A. Robock, C. Ammann, L. Oman, D. Shindell, and G. Stenchikov. Can volcanic eruptions produce ice ages or Mass extinctions? *AGU Fall Meeting Abstracts*, pages G6+, December 2006.
212. D. J. Roddy, E. M. Shoemaker, and R. R. Anderson. The Manson impact crater: Estimation of the energy of formation, possible size of the impacting asteroid or comet, and ejecta volume and mass. In *Lunar and Planetary Institute Conference Abstracts*, pages 1211–1212, March 1993.
213. G. Rohen, C. von Savigny, M. Sinnhuber, E. J. Llewellyn, J. W. Kaiser, C. H. Jackman, M.-B. Kallenrode, J. Schröter, K.-U. Eichmann, H. Bovensmann, and J. P. Burrows. Ozone depletion during the solar proton events of October/November 2003 as seen by SCIAMACHY. *Journal of Geophysical Research (Space Physics)*, 110(9):9, August 2005.
214. V. S. Safronov and E. V. Zvjagina. Relative sizes of the largest bodies during the accumulation of planets. *Icarus*, 10:109, January 1969.
215. C. Sagan and C. Chyba. The early faint sun paradox: Organic shielding of ultraviolet-labile greenhouse gases. *Science*, 276:1217–1221, 1997.
216. C. Sagan and E. E. Salpeter. Particles, environments, and possible ecologies in the Jovian atmosphere. *Astrophysical Journal Supplement*, 32:737–755, October 1976.
217. Z. Sándor, Á. Süli, B. Érdi, E. Pilat-Lohinger, and R. Dvorak. A stability catalogue of the habitable zones in extrasolar planetary systems. *Monthly Notices*, 375:1495–1502, March 2007.
218. J. W. Schopf. Microfossils of the early archean apex chert: New evidence of the antiquity of life. *Science*, 260:640–646, April 1993.
219. D. Schulze-Makuch, D. H. Grinspoon, O. Abbas, L. N. Irwin, and M. A. Bullock. A sulfur-based survival strategy for putative phototrophic life in the venusian atmosphere. *Astrobiology*, 4:11–18, March 2004.

220. B. C. Schuraytz, D. J. Lindstrom, L. E. Marin, R. R. Martinez, D. W. Mittlefehldt, V. L. Sharpton, and S. J. Wentworth. Iridium metal in chicxulub impact melt: Forensic chemistry on the K-T smoking gun. *Science*, 271:1573–1576, March 1996.
221. Z. Sekanina. The Tunguska event – No cometary signature in evidence. *The Astronomical Journal*, 88:1382–1413, September 1983.
222. Z. Sekanina. The problem of split comets revisited. *Astronomy and Astrophysics*, 318:L5–L8, February 1997.
223. M. A. Sephton and O. Botta. Extraterrestrial organic matter and the detection of life. *Space Science Reviews*, pages 103–+, May 2007.
224. T. V. Shabanova. Evidence for a planet around the pulsar PSR B0329+54. *Astrophysical Journal*, 453:779, November 1995.
225. P. M. Sheehan and D.E. Fastovsky. Major extinctions of land-dwelling vertebrates at the Cretaceous-Tertiary boundary, eastern Montana. *Geology*, 20:556-560.
226. E. M. Shoemaker. Geological and astronomical evidence for comet impact and comet showers during the last 100 million years. *LPI Contributions*, 765:199, 1991.
227. E. M. Shoemaker and G. A. Izett. K/T boundary stratigraphy: Evidence for multiple impacts and a possible comet stream. *LPI Contributions*, 790:66, 1992.
228. N. H. Sleep and K. Zahnle. Refugia from asteroid impacts on early Mars and the early Earth. *Journal of Geophysical Research*, 103:28529–28544, December 1998.
229. N. Smith, W. Li, R. J. Foley, J. C. Wheeler, D. Pooley, R. Chornock, A. V. Filippenko, J. M. Silverman, R. Quimby, J. S. Bloom, and C. Hansen. SN 2006gy: Discovery of the most Luminous Supernova ever recorded, powered by the death of an extremely massive star like η Carinae. *Astrophysical Journal*, 666:1116–1128, September 2007.
230. L. E. Snyder, F. J. Lovas, J. M. Hollis, D. N. Friedel, P. R. Jewell, A. Remijan, V. V. Ilyushin, E. A. Alekseev, and S. F. Dyubko. A rigorous attempt to verify interstellar glycine. *Astrophysical Journal*, 619:914–930, February 2005.
231. F. Sohl and H. Hussmann. Subsurface Oceans on the Saturnian Satellites. In *Bulletin of the American Astronomical Society, volume 37 of Bulletin of the American Astronomical Society*, page 1569, December 2005.
232. L. S. Sparke and J.S. Gallagher III. *Galaxies in the Uuniverse: An Introduction*, Cambridge , Cambridge University Press, 2000.
233. T. Spohn and G. Schubert. Oceans in the icy Galilean satellites of Jupiter? *Icarus*, 161:456–467, February 2003.
234. D. I. Steel. Collisions in the solar system. V - Terrestrial impact probabilities for parabolic comets. *Monthly Notices*, 264:813, October 1993.
235. T. F. Stepinski and P. Valageas. Global evolution of solid matter in turbulent protoplanetary disks. II. Development of icy planetesimals. *Astronomy and Astrophysics*, 319:1007–1019, March 1997.
236. K. R. Stern, S. Jansky and J.E. Blidlack. *Introductory Plant Biology*, McGraw Hill, 2003.
237. D. J. Stevenson. Life-sustaining planets in interstellar space? *Nature*, 400:32–+, July 1999.
238. C. R. Stoker and M. A. Bullock. Organic degradation under simulated Martian conditions. *Journal of Geophysical, Research*, 102:10881–10888, May 1997.
239. R. B. Stothers. Terrestrial record of the solar system's oscillation about the galactic plane. *Nature*, 317:338–341, September 1985.
240. C. Struck. *Galaxy Collisions - Dawn of a New Era*, page 115. Astrophysics Update 2, 2006.
241. Z. Suo, R. Avci, M. Higby, and M. Deliorman. Porphyrin as an ideal biomarker in the search for extraterrestrial life. *Astrobiology*, 7:605–615, August 2007.
242. T. Takata, T. J. Ahrens, and A. W. Harris. Fragment and Progenitor Energy of Comet Shoemaker-Levy 9 and Frequency of Such Impact Events. In *Lunar and Planetary Institute Conference Abstracts*, volume 26 of *Lunar and Planetary Institute Conference Abstracts*, page 1393, March 1995.
243. P. Thébault, J. C. Augereau, and H. Beust. Dust production from collisions in extrasolar planetary systems. The inner beta Pictoris disc. *Astronomy and Astrophysics*, 408:775–788, September 2003.

244. B. C. Thomas, C. H. Jackman, A. L. Melott, C. M. Laird, R. S. Stolarski, N. Gehrels, J. K. Cannizzo, and D. P. Hogan. Terrestrial ozone depletion due to a Milky Way gamma-ray burst. *Astrophysical Journal Letters*, 622:L153–L156, April 2005.
245. B. C. Thomas, A. L. Melott, B. D. Fields, and B. J. Anthony-Twarog. Superluminous supernovae: No threat from Eta Carinae. *ArXiv e-prints*, 705, May 2007.
246. H. B. Throop and J. Bally. Can photoevaporation trigger planetesimal formation? *Astrophysical Journal Letters*, 623:L149–L152, April 2005.
247. B. T. Tsurutani, A. J. Mannucci, B. Iijima, F. L. Guarnieri, W. D. Gonzalez, D. L. Judge, P. Gangopadhyay, and J. Pap. The extreme Halloween 2003 solar flares (and Bastille Day, 2000 Flare), ICMEs, and resultant extreme ionospheric effects: A review. *Advances in Space Research*, 37:1583–1588, 2006.
248. S. Turck-Chieze, S. Cahen, M. Casse, and C. Doom. Revisiting the standard solar model. *Astrophysical Journal*, 335:415–424, December 1988.
249. P. Ulmschneider. *Intelligent Life in The Universe*, 2nd Ed., Springer, 2006.
250. J.-P. Valet and L. Meynadier. Geomagnetic field intensity and reversals during the past four million years. *Nature*, 366:234–238, November 1993.
251. A. M. Vázquez, Hanslmeier. *Ultraviolet Radiation in the Solar System.* Springer, *Astrophysics and Space Science Library*, 331, 2005.
252. W. von Bloh, C. Bounama, M. Cuntz, and S. Franck. Habitability of Super-Earths: Gliese 581c and 581d. *ArXiv e-prints*, 712, December 2007.
253. W. von Bloh, C. Bounama, and S. Franck. *Habitable Zones in Extrasolar Planetary Systems*, pages 255–+. Extrasolar Planets. Formation, Detection and Dynamics, Edited by Rudolf Dvorak. ISBN: 978-3-527-40671-5 Wiley-VCH, 2007, p. 255., 2007.
254. B. Vukotic and M. M. Cirkovic. On the timescale forcing in astrobiology. *Serbian Astronomical Journal*, 175:45–50, December 2007.
255. G. Waechtershaeuser. Before enzymes and templates: Theory of surface metabolism. *Microbiological Review.*, 54:452, December 1988.
256. G. A. H. Walker, J. M. Matthews, R. Kuschnig, J. F. Rowe, D. B. Guenther, A. F. J. Moffat, S. Rucinski, D. Sasselov, S. Seager, E. Shkolnik, and W. W. Weiss. Precise photometry of 51 Peg systems with MOST. In L. Arnold, F. Bouchy, and C. Moutou, editors, *Tenth Anniversary of 51 Peg-b: Status of and prospects for hot Jupiter studies*, pages 267–273, February 2006.
257. Z. Wang, D. Chakrabarty, and D. L. Kaplan. A debris disk around an isolated young neutron star. *Nature*, 440:772–775, April 2006.
258. W. R. Ward. Comments on the long-term stability of the earth's obliquity. *Icarus*, 50:444–448, June 1982.
259. W. R. Ward. Protoplanet migration by nebula tides. *Icarus*, 126:261–281, April 1997.
260. A. J. Watson and J. E. Lovelock. Biological homeostasis of the global environment: the parable of Daisyworld. *Tellus Series B Chemical and Physical Meteorology B*, 35:284, 1983.
261. S. Webb. Where is Everybody? Fifty Solutions to the Fermi Paradox and the Problem of Extraterrestrial Life, Copernicus books, New York, Springer, 2002.
262. A. J. Weinberger. On the binary nature of dust-encircled BD +20 307. *Astrophysical Journal Letters*, 679:L41–L44, May 2008.
263. P. R. Weissman. The Cometary Impactor Flux at the Earth. In *IAU Symposium*, August 2006.
264. G. W. Wetherill. Formation of the terrestrial planets. *Astronomy and Astrophysics*, 18:77–113, 1980.
265. G. W. Wetherill. Ratio of Asteroidal Impact Rates on Mars and Earth. In *Lunar and Planetary Institute Conference Abstracts, volume 18 of Lunar and Planetary Institution of. Technical Report*, page 1076, March 1987.
266. W. A. Wheaton, M. P. Ulmer, W. A. Baity, D. W. Datlowe, M. J. Elcan, L. E. Peterson, R. W. Klebesadel, I. B. Strong, T. L. Cline, and U. D. Desai. The direction and spectral variability of a cosmic gamma-ray burst. *Astrophysical Journal Letters*, 185:L57+, October 1973.
267. F. L. Whipple. On the structure of the cometary nucleus. 65:503, 1960.

268. D. P. Whitmire and A. A. Jackson. Are periodic mass extinctions driven by a distant solar companion? *Nature*, 308:713–715, April 1984.
269. D. M. Williams, J. F. Kasting, and R. A. Wade. Habitable moons around extrasolar giant planets. *Nature*, 385:234–236, January 1997.
270. A. Wolszczan and D. A. Frail. A planetary system around the millisecond pulsar PSR1257 + 12. *Nature*, 355:145–147, January 1992.
271. D. K. Yeomans. The Impact of Comets and Asteroids Upon the Earth. In *Bulletin of the American Astronomical Society*, page 1038, September 1998.
272. M. R. Zapatero Osorio, V. J. S. Béjar, E. L. Martín, R. Rebolo, D. Barrado y Navascués, C. A. L. Bailer-Jones, and R. Mundt. Discovery of young, isolated planetary mass objects in the σ Orionis star cluster. *Science*, 290:103–107, October 2000.
273. M. R. Zapatero Osorio, V. J. S. Béjar, E. L. Martín, R. Rebolo, D. Barrado y Navascués, R. Mundt, J. Eislöffel, and J. A. Caballero. A Methane, isolated, planetary-mass object in Orion. *Astrophysical Journal*, 578:536–542, October 2002.
274. S. Zhang. Molecular self-assembly: Another brick in the wall. *Nature*, 1:169, 2006.
275. B. Zhang and S. Sigurdsson. Electromagnetic signals from planetary collisions. *Astrophysical Journal Letters*, 596:L95–L98, October 2003.
276. H. Zinnecker, S. Correia, W. Brandner, S. Friedrich, and M. McCaughrean. Search for giant extrasolar planets around white dwarfs: direct imaging with NICMOS/HST and NACO/VLT. In C. Aime and F. Vakili, editors, *IAU Colloq. 200: Direct Imaging of Exoplanets: Science and Techniques*, pages 19–24, 2006.

Index