The Solar System

7

The solar system consists of a central star, called *the Sun*, *planets* orbiting the Sun and several smaller objects. By the solar system we mean here and in the next chapter the system around our own Sun. Also other stars have similar systems. They will be discussed later in Chap. 22. This chapter deals with general properties of the solar system. Individual objects will be discussed in the next chapter.

Research of the solar system has evolved rapidly since the 1960's when space probes have made it possible to study planets from short distances. Many methods used in geosciences are nowadays applied also in planetary studies. Landers have been sent to the Moon, Venus, Mars, and several smaller bodies.

The most convenient way to describe distances in the solar system is to use *astronomical units* (au), the mean distance of the Sun and Earth. One au is $= 1.495\,978\,70 \times 10^{11}$ m (see Sect. 6.5). The distance to the nearest star, *Proxima Centauri* is over 270,000 au.

Actually one au is the semimajor axis of a massless planet whose orbital period equals that of the Earth. Since also the mass of the Earth affects its motion, the semimajor axis of the Earth is slightly larger than one au (6.32). The small difference is, however, important only in precise calculations.

7.1 Classification of Objects

In addition to the Sun and Moon five objects moving relative to the stars were already known in the antiquity: Mercury, Venus, Mars, Jupiter and Saturn. They were called planets from the Greek word meaning a wanderer. At that time also the Sun and Moon were considered planets, and the names of these seven objects are still reflected in the names of the days of the week.

After the invention of the telescope three more planets were found: Uranus, Neptune and Pluto. When observational instruments and methods have evolved, more objects orbiting behind Neptune and of the same size as Pluto have been found. Since there was no unambiguous definition of a planet, some of these new objects should have been called planets. Therefore the *International Astronomical Union* (IAU) in its General Assembly in 2006 defined three distinct categories to clarify the situation.

According to the new definition, an object is a *planet* if it satisfies the following three conditions:

(1) It orbits the Sun.
(2) It has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a hydrostatic equilibrium (nearly round) shape.
(3) Its perturbations have cleared away other objects in the neighbourhood of its orbit.

If a body satisfies the conditions (1) and (2) but not (3), it is a *dwarf planet*. Since Pluto does not fulfil the last requirement its status was reduced to a dwarf planet.

All other objects orbiting the Sun shall be referred to collectively as *Small Solar System Bodies*. These include most of the asteroids, Trans-

H. Karttunen et al. (eds.), *Fundamental Astronomy*, DOI 10.1007/978-3-662-53045-0_7

Neptunian Objects, comets, and other small bodies. If the central body is not the Sun the object is a *moon* or a *satellite* independently of its properties.

A *satellite* is a body which orbits the primary body so that the centre of mass (barycentre) is inside the primary. If this is not the case, then the system is called a *binary system*. For example, in the case of the Earth and Moon the barycentre of the system is inside the Earth, and the Moon is Earth's satellite. In the Pluto-Charon system the centre of mass is outside Pluto, and therefore they are called a binary system.

The rules (2) and (3) are somewhat problematic. How can the shape of a very distant object be established? How wide an area a planet should clean to be a planet, and how clean that region should be?

Also the definition of a satellite is slightly problematic, since the largest planets are surrounded by ring systems consisting of small particles. What is the distinction between the largest ring particles and smallest satellites? In practice, an object could be considered a moon if it has been observed as a separate body whose orbital elements have been determined. In any case the question about the exact number of satellites is no more meaningful, since more and more will be found when observations become more accurate. More interesting questions are e.g. the statistical properties of the moons.

Other unresolved questions concern the classification of the minor bodies and exoplanets. Possibly the definitions need some fine-tuning in the future.

According the current definition there are now eight planets in the solar system: *Mercury*, *Venus*, *Earth*, *Mars*, *Jupiter*, *Saturn*, *Uranus*, and *Neptune*. It looks quite improbable that more planets will be found.

Currently there are five known dwarf planets: *Pluto* that was earlier a planet, the first asteroid *Ceres* and *Haumea*, *Makemake* and *Eris* found in 2004–2005 orbiting beyond Pluto. Similar objects may well be found in the future. The problem is to check whether these distant bodies satisfy the conditions (2) and (3).

The solar system contains a vast number of different small bodies. Traditionally, they have been divided into three categories: *asteroids*, *comets* and *meteoroids*. Their differences and properties are discussed in more detail in Sects. 8.10–8.14.

The planets from Mercury to Saturn are bright and well visible with a naked eye. Records of them are found even in the most ancient written documents. Uranus and Neptune can be seen with a pair of binoculars. In addition to the bright planets, only the brightest comets are visible with a naked eye.

Gravitation controls the motion of the solar system bodies. The planetary orbits around the Sun (Fig. 7.1) are almost coplanar ellipses which deviate only slightly from circles. Mercury, the innermost planet, has the most eccentric orbit. The orbital planes of *asteroids*, minor bodies that circle the Sun mainly between the orbits of Mars and Jupiter, are often more tilted than the planes of the planetary orbits. Asteroids and distant Trans-Neptunian Objects revolve in the same direction as the major planets; comets, however, may move in the opposite direction. Cometary orbits can be very elongated, even hyperbolic. Most of the satellites circle their parent planets in the same direction as the planet moves around the Sun. Only the motions of the smallest particles, gas and dust are affected by the *solar wind*, *radiation pressure* and *magnetic fields*.

The planets can be divided into inferior and superior planets depending on their orbits. Mercury and Venus are inferior planets as seen from the Earth, and planets from Mars to Neptune are superior planets.

Depending on their physical properties the planets can be divided into different groups (Fig. 7.2). Mercury, Venus, Earth, and Mars are called *terrestrial* (Earth-like) planets; they have a solid surface, are of almost equal size (diameters from 5000 to 12,000 km), and have quite a high mean density (4000–5000 $\mathrm{kg\,m^{-3}}$; the density of water is 1000 $\mathrm{kg\,m^{-3}}$). The planets from Jupiter to Neptune are called *Jovian* (Jupiter-like) or *giant planets*. The densities of the giant planets are about 1000–2000 $\mathrm{kg\,m^{-3}}$, and most of their volume is liquid. Diameters are several times greater than those of the terrestrial planets.

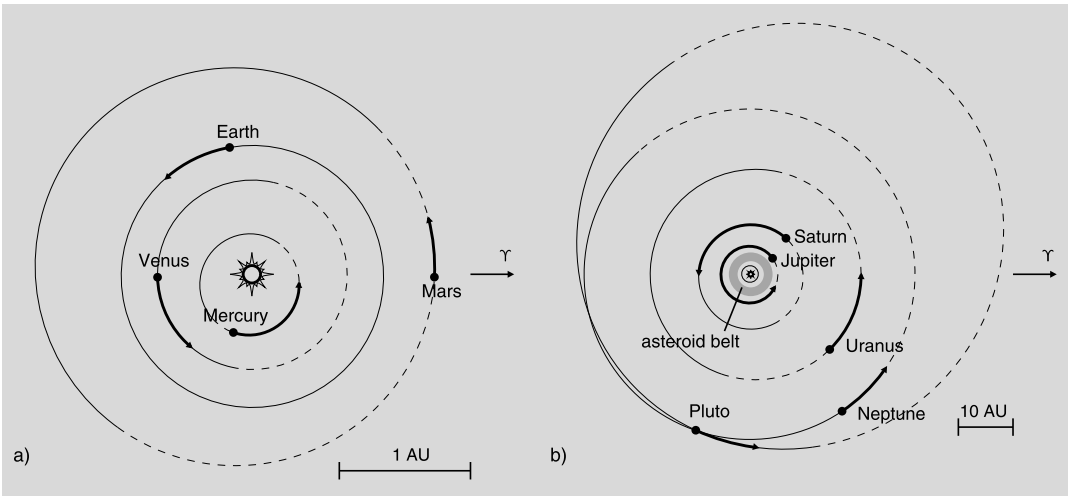


Fig. 7.1 (**a**) Planetary orbits from Mercury to Mars. *The dashed line* represents the part of the orbit below the ecliptic; *the arrows* show the distances travelled by the planets during one month (January 2000). (**b**) Planets from Jupiter to Neptune and the dwarf planet Pluto. *The arrows* indicate the distances travelled by the planets during the 10 year interval 2000–2010

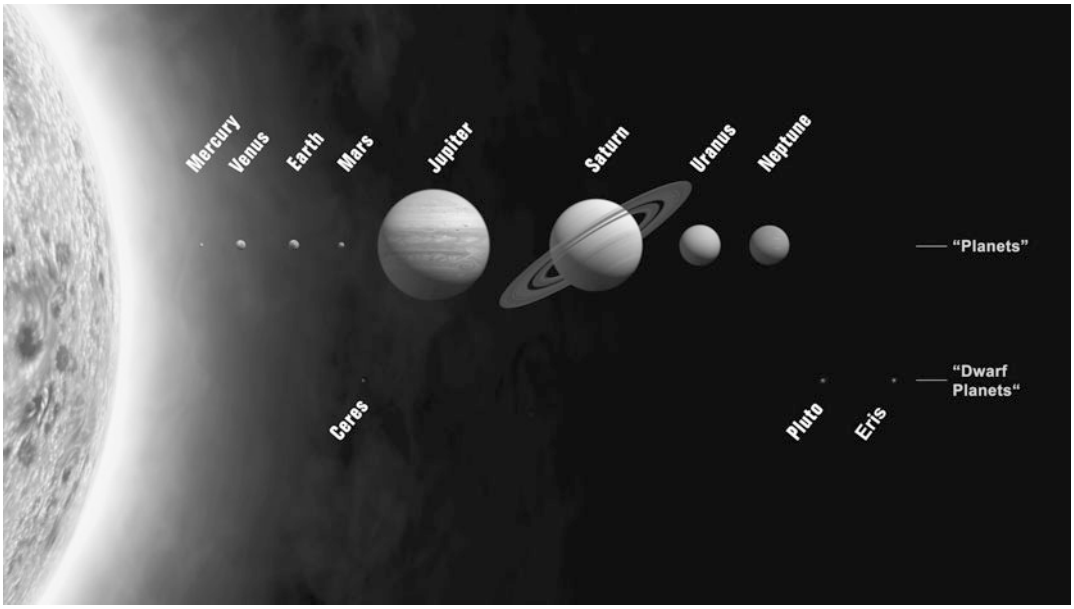


Fig. 7.2 Major planets from Mercury to Neptune. Four innermost planets are called terrestrial planets and four outermost ones are giant planets. Three dwarf planets are also shown. Relative size of the Sun is shown *at left*. Planetary distances to the Sun are not in scale. (The International Astronomical Union/Martin Kornmesser)

7.2 Planetary Configurations

To the naked eye planets look starlike dots. However, their slow motion with respect to the stars reveals that they are bodies of our solar system. The *apparent motions* of the planets look quite complicated, partly because they reflect the motion of the Earth around the Sun (Figs. 7.3 and 7.4).

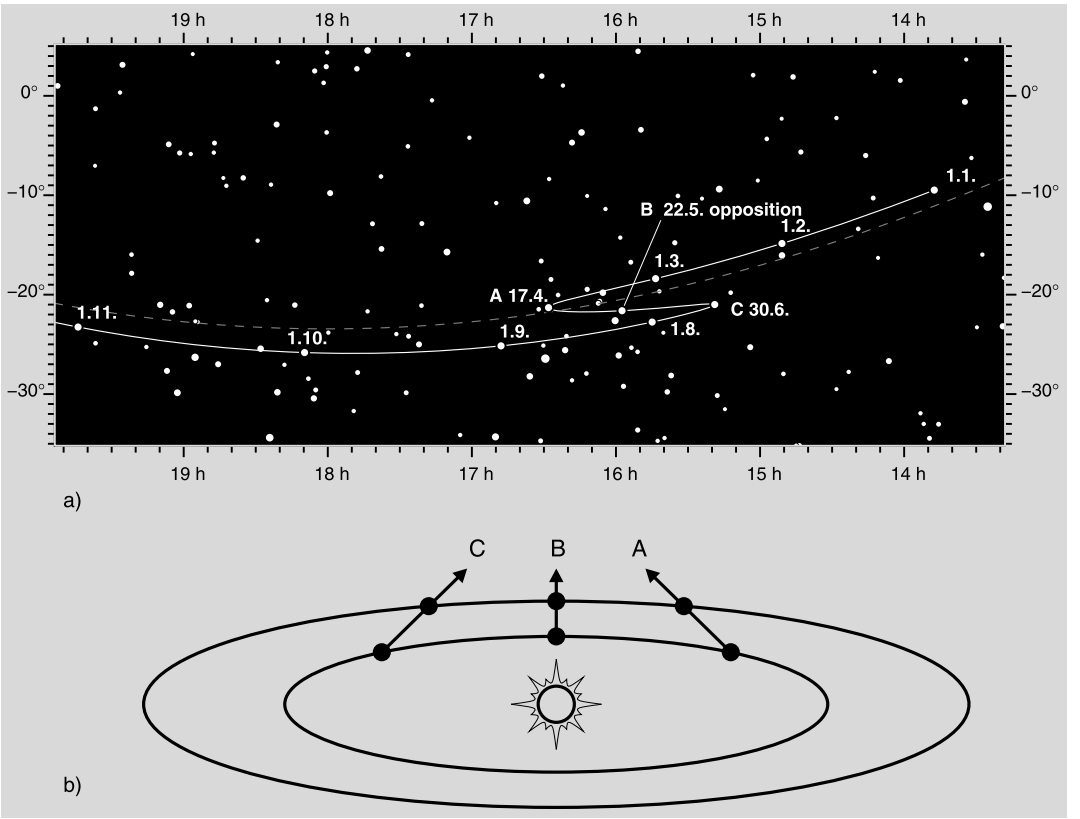


Fig. 7.3 (**a**) Apparent motion of Mars during the 2016 opposition. Usually Mars moves in the prograde direction (counterclockwise with respect to the stars), but over a month before and after the opposition the motion is retrograde. (**b**) Relative positions of the Earth and Mars. The projection of the Earth–Mars direction on the infinitely distant celestial sphere results in (**a**)

Normally the planets move eastward (*direct motion*, counterclockwise as seen from the Northern hemisphere) when compared with the stars. When the Earth passes a superior planet, the motion of the planet reverses to the opposite or *retrograde* direction. After a few weeks of retrograde motion, the direction is changed again, and the planet continues in the original direction. It is quite understandable that the ancient astronomers had great difficulties in explaining and modelling such complicated turns and loops. Figure 7.5 explains some basic planetary configurations.

A superior planet (planet outside the orbit of the Earth) is said to be in *opposition* when it is exactly opposite the Sun, i.e. when the Earth is between the planet and the Sun. When the planet is behind the Sun, it is in *conjunction*. In practise, the planet may not be exactly opposite or behind the Sun because the orbits of the planet and the Earth are not in the same plane. In astronomical almanacs oppositions and conjunctions are defined in terms of ecliptic longitudes. The longitudes of a body and the Sun differ by 180° at the moment of opposition; in conjunction the longitudes are equal. However, the right ascension is used if the other body is not the Sun. Those points where the apparent motion of a planet changes its direction are called *stationary points*. Opposition occurs in the middle of the retrograde loop.

Inferior planets (Mercury and Venus) are never in opposition. The configuration when the planets is between the Earth and the Sun is called *inferior conjunction*. The conjunction corresponding to that of a superior planet is called an *upper*

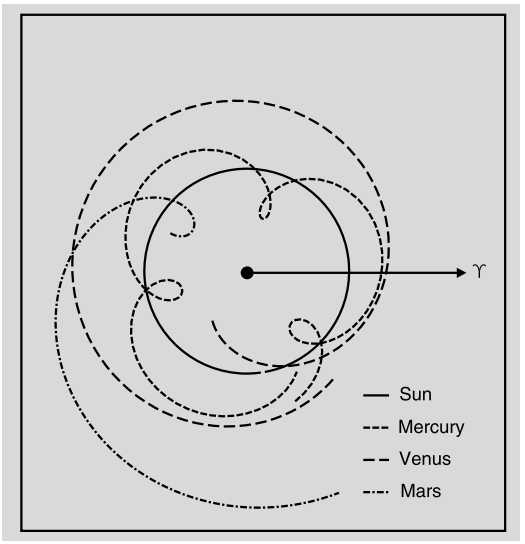


Fig. 7.4 Apparent motions of the Sun, Mercury, Venus and Mars in 1995 in the geocentric frame as seen from the North pole of the ecliptic

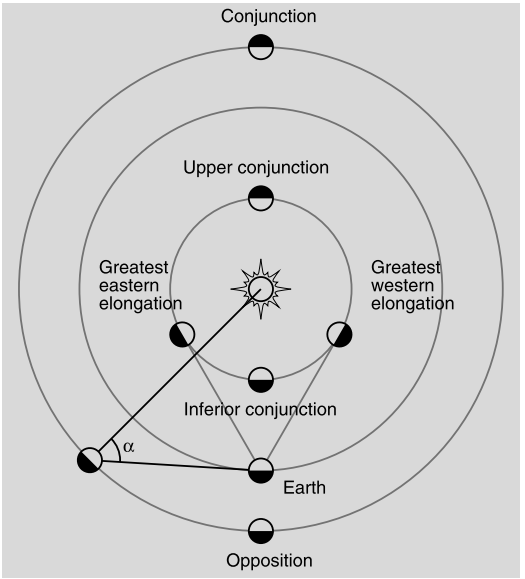


Fig. 7.5 Planetary configurations. The angle α (Sun–object–Earth) is the phase angle and ε (Sun–Earth–object) is the elongation

conjunction or *superior conjunction*. The maximum (eastern or western) *elongation*, i.e. the angular distance of the planet from the Sun is 28° for Mercury and 47° for Venus. Elongations are called eastern or western, depending on which side of the Sun the planet is seen. The planet is an "evening star" and sets after the Sun when it is in eastern elongation; in western elongation the planet is seen in the morning sky as a "morning star".

The *synodic period* is the time interval between two successive events (e.g. oppositions). The period which we used in the previous chapters is the *sidereal period*, the true time of revolution around the Sun, unique for each object. The synodic period depends on the difference of the sidereal periods of the two bodies.

Let the sidereal periods of two planets be P_1 and P_2 (assume that $P_1 < P_2$). Their mean angular velocities (*mean motions*) are $2\pi/P_1$ and $2\pi/P_2$. After one synodic period $P_{1,2}$, the inner planet has made one full revolution more than the outer planet. Thus the angles travelled by the planets are related by

$$P_{1,2}\frac{2\pi}{P_1} = 2\pi + P_{1,2}\frac{2\pi}{P_2},$$

or

$$\frac{1}{P_{1,2}} = \frac{1}{P_1} - \frac{1}{P_2}. \tag{7.1}$$

Since Mercury and Venus move inside the orbit of the Earth, they show similar *phases* as the Moon. The angle Sun–planet–Earth is called the *phase angle*, often denoted by the Greek letter α. The fraction of the planetary surface seen illuminated depends on the phase angle. For Mercury and Venus the phase angle can have any value between 0° and 180°. This means that we can see "full Venus" (when it is behind the Sun), "half Venus", and so on.

The phase angle range of the superior planets is more limited. For Mars the maximum phase angle is 41°, for Jupiter 11°, and for Neptune only 2°.

7.3 Orbit of the Earth and Visibility of the Sun

The rotation of the Earth and its motion around the Sun have been used as the basis of time reckoning since prehistoric times (although their true

nature was not known then). However, these motions are not free from perturbations. Although our modern time system is based on atomic clocks, the ordinary time we use in our clocks is adjusted to the rotation of the Earth by leap seconds (Sect. 2.14).

The *sidereal year* is the real orbital period of the Earth around the Sun. After one sidereal year, the Sun is seen at the same position relative to the stars. The length of the sidereal year is 365.256363051 days of 86,400 SI seconds at the epoch J2000.0 = 2000 January 1 12:00:00 TT.

We noted earlier that, owing to precession, the direction of the vernal equinox moves along the ecliptic at about 50″ per year. This means that the Sun returns to the vernal equinox before one complete sidereal year has elapsed. This time interval, called the *tropical year*, is 365.24218967 days.

A third definition of the year is based on the perihelion passages of the Earth. Planetary perturbations cause a gradual change in the direction of the Earth's perihelion. The time interval between two successie perihelion passages is called the *anomalistic year*, the length of which is 365.259635864 days, a little longer than the sidereal year. It takes about 21,000 years for the perihelion to revolve 360° relative to the vernal equinox.

The equator of the Earth is tilted about 23.4° with respect to the ecliptic. Owing to perturbations, this angle changes with time. If periodic terms are neglected, the *obliquity of the ecliptic* ε can be calculated as in Box 2.1. The obliquity varies between 22.1° and 24.5° with a 41,000 year periodicity. At present the tilt is decreasing. There are also small short term variations, like the *nutation*.

The declination of the Sun varies between $-\varepsilon$ and $+\varepsilon$ during the year. At any given time, the Sun is seen at zenith from one point on the surface of the Earth. The latitude of this point is the same as the declination of the Sun. At the latitudes $-\varepsilon$ (the *Tropic of Capricorn*) and $+\varepsilon$ (the *Tropic of Cancer*), the Sun is seen at zenith once every year, and between these latitudes twice a year. The Sun crosses the equator at vernal and autumnal equinox.

In the Northern hemisphere the Sun will not set if the latitude is greater than $90° - \delta$, where δ is the declination of the Sun. The southernmost latitude where the *midnight Sun* can be seen is thus $90° - \varepsilon = 66.6°$. This is called the *Arctic Circle*. (The same holds true in the Southern hemisphere.) The Arctic Circle is the southernmost place where the Sun is (in theory) below the horizon during the whole day at the winter solstice. The sunless time lasts longer and longer when one goes north (south in the Southern hemisphere). At the poles, day and night last half a year each. In practise, refraction and location of the observing site will have a large influence on the visibility of the midnight Sun and the number of sunless days. Because refraction raises objects seen at the horizon, the midnight Sun can be seen a little further south than at the Arctic Circle. For the same reason the Sun can be seen simultaneously at both poles around the time of vernal and autumnal equinox. (See also Sect. 2.6.)

The eccentricity of the Earth's orbit is about 0.0167. The distance from the Sun varies between 147–152 million km. The flux density of solar radiation varies somewhat at different parts of the Earth's orbit, but this has practically no effect on the seasons. In fact the Earth is at perihelion in the beginning of January, in the middle of the northern hemisphere's winter.

The *seasons* are due to the obliquity of the ecliptic which affects the energy received from the Sun depends in three different ways. First the flux per unit area is proportional to $\sin a$, where a is the altitude of the Sun. In summer the altitude can have greater values than in winter, giving more energy per unit area. Another effect is due to the atmosphere: when the Sun is near the horizon, the radiation must penetrate thick atmospheric layers. This means large extinction and less radiation at the surface. The third factor is the length of the time the Sun is above the horizon. This is important at high latitudes, where the low altitude of the Sun is compensated by the long daylight time in summer. This last effect is neglected in many textbooks written by authors not living this far north. These effects are discussed in detail in Example 7.2.

There are also long-term variations in the annual Solar flux. Serbian geophysicist Milutin Milanković (1879–1958) published in the 1930's

and 1940's his theory of ice ages. During the last 2–3 million years, large ice ages have recurred approximately every 100,000 years. He proposed that variations of the Earth's orbit cause long-term periodic climate change, now known as *Milanković cycles*. Milanković claimed that the cycles in eccentricity, direction of the perihelion, obliquity, and precession result in 100,000 year ice age cycle. The cycle of precession is 26,000 years, direction of the perihelion relative to the equinoxes is 22,000 years, and the obliquity of the ecliptic has a 41,000 year cycle. Changes in orbital eccentricity are not fully periodic but some periods above 100,000 years can be found. The eccentricity varies between 0.005–0.058 and is currently 0.0167.

The annual incoming Solar flux varies with these orbital changes and the effect is largest at high latitudes. If, for example, the eccentricity is high, and the Earth is near the aphelion during the hemisphere's winter, then winters are long and cold and summers are short. However, the theory is controversial, orbital forcing on the climate change is not well understood, and probably not enough to trigger glaciation. There exist also positive feedback loops, like the effect of low albedo of snow and ice. It means that ice reflects more radiation back into space, thus cooling the climate. The system is highly chaotic so that even minor changes in the primary conditions will result in large differences in the outcome. There are also other effects causing climate change, including emerging gases from large lava flows and eruptions of volcanos and, nowadays, anthropogenic reasons.

The future is also uncertain. Some theories predict that the warm period will continue next 50,000 years, whereas others conclude that the climate is already cooling. Anthropogenic reasons, like ever increasing fraction of green house gases, e.g. carbon dioxide, will change the short-term predictions.

7.4 The Orbit of the Moon

The Earth's satellite, *the Moon*, circles the Earth counterclockwise. One revolution, the *sidereal month*, takes about 27.322 days. In practise, a more important period is the *synodic month*, the duration of the Lunar phases (e.g. from full moon to full moon). In the course of one sidereal month the Earth has travelled almost $1/12$ of its orbit around the Sun. The Moon still has about $1/12$ of its orbit to go before the Earth–Moon–Sun configuration is again the same. This takes about 2 days, so the phases of the Moon are repeated every 29 days. More exactly, the length of the synodic month is 29.531 days.

The *new moon* is that instant when the Moon is in conjunction with the Sun. Almanacs define the phases of the Moon in terms of ecliptic longitudes; the longitudes of the new moon and the Sun are equal. Usually the new moon is slightly north or south of the Sun because the lunar orbit is tilted 5° with respect to the ecliptic.

About 2 days after the new moon, the waxing crescent moon can be seen in the western evening sky. About 1 week after the new moon, the *first quarter* follows, when the longitudes of the Moon and the Sun differ by 90°. The right half of the Moon is seen lit (left half when seen from the Southern hemisphere). The *full moon* appears a fortnight after the new moon, and 1 week after this the *last quarter*. Finally the waning crescent moon disappears in the glory of the morning sky.

The orbit of the Moon is approximately elliptic. The length of the semimajor axis is 384,400 km and the eccentricity 0.055. Owing to perturbations caused mainly by the Sun, the orbital elements vary with time. The minimum distance of the Moon from the centre of the Earth is 356,400 km, and the maximum distance 406,700 km. This range is larger than the one calculated from the semimajor axis and the eccentricity. The apparent angular diameter is in the range $29.4'$–$33.5'$.

The rotation time of the Moon is equal to the sidereal month, so the same side of the Moon always faces the Earth. Such *synchronous rotation* is common among the satellites of the solar system: almost all large moons rotate synchronously.

The orbital speed of the Moon varies according to Kepler's second law. The rotation period, however, remains constant. This means that, at

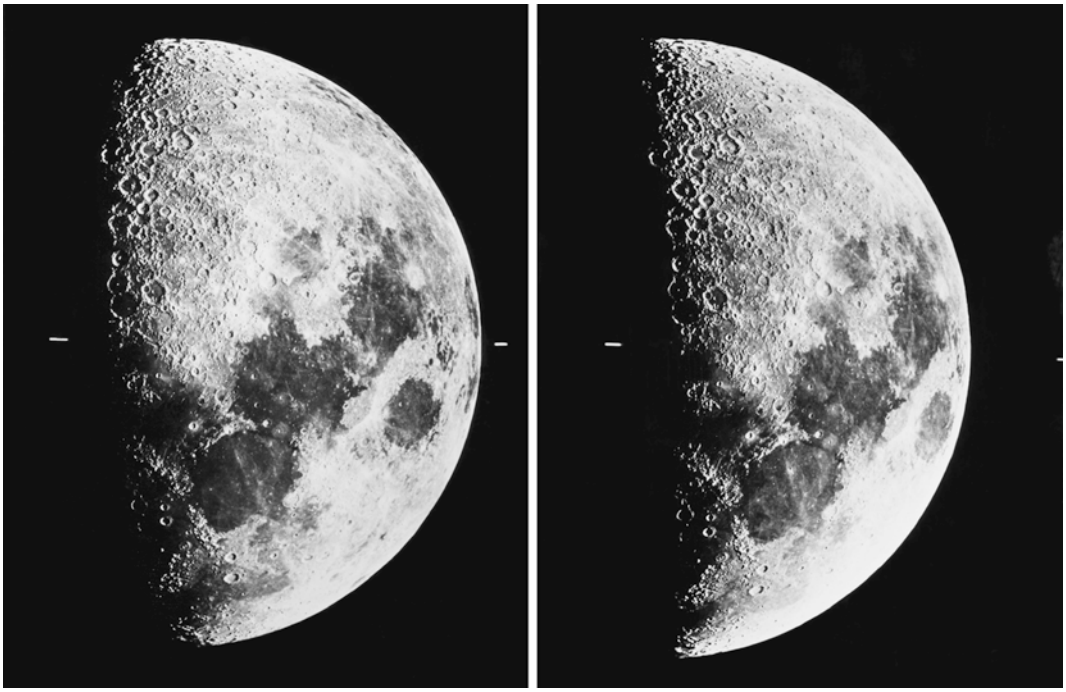


Fig. 7.6 Librations of the Moon can be seen in this pair of photographs taken when the Moon was close to the perigee and the apogee, respectively. (Helsinki University Observatory)

different phases of the lunar orbit, we can see slightly different parts of the surface. When the Moon is close to its perigee, its speed is greater than average (and thus greater than the mean rotation rate), and we can see more of the right-hand edge of the Moon's limb (as seen from the Northern hemisphere). Correspondingly, at the apogee we see "behind" the left edge. Owing to this *libration*, a total of 59 % of the surface area can be seen from the Earth (Fig. 7.6). The libration is quite easy to see if one follows some detail at the edge of the lunar limb.

There are also two other factors causing libration. When the Moon is rising from the East, we can see a little behind the right edge, and when the Moon is setting, behind the left edge. The third effect is the latitudinal libration. Since the orbit of the Moon is not quite in the equatorial plane, we can see behind the northern or southern pole depending on the position of the Moon.

The orbital plane of the Moon is tilted only about 5° to the ecliptic. Therefore the Moon is always close to the ecliptic just like the Sun and the planets. However, the orbital plane changes gradually with time, owing mainly to the perturbations caused by the Earth and the Sun. These perturbations cause the nodal line (the intersection of the plane of the ecliptic and the orbital plane of the Moon) to make one full revolution in 18.6 years. We have already encountered the same period in the nutation. When the ascending node of the lunar orbit is close to the vernal equinox, the Moon can be $23.4° + 5° = 28.4°$ north or south of the equator. When the descending node is close to the vernal equinox, the zone where the Moon can be found extends only $23.4° - 5° = 18.4°$ north or south of the equator.

The *nodical* or *draconic month* is the time in which the Moon moves from one ascending node back to the next one. Because the line of nodes is rotating, the nodical month is 3 hours shorter than the sidereal month, i.e. 27.212 days. The orbital ellipse itself also precesses slowly. The orbital period from perigee to perigee, the *anomalistic month*, is 5.5 h longer than the sidereal month, or about 27.555 days.

Gravitational differences caused by the Moon and the Sun on different parts of the Earth's surface give rise to the *tides*. Gravitation is greatest at the sub-lunar point and smallest at the opposite side of the Earth. At these points, the surface of the seas is highest (high tide, *flood*). About 6 h after flood, the surface is lowest (low tide, *ebb*). The tide generated by the Sun is less than half of the lunar tide. When the Sun and the Moon are in the same direction with respect to the Earth (new moon) or opposite each other (full moon), the tidal effect reaches its maximum; this is called *spring tide*.

The sea level typically varies 1 m, but in some narrow straits, the difference can be as great as 15 m. Due to the irregular shape of the oceans, the true pattern of the oceanic tide is very complicated.

The solid surface of the Earth also suffers tidal effects, but the amplitude is much smaller, about 30 cm.

Tides generate friction, which dissipates the rotational and orbital kinetic energy of the Earth–Moon system. This energy loss induces some changes in the system. First, the rotation of the Earth slows down until the Earth also rotates synchronously, i.e. the same side of Earth will always face the Moon. Secondly, the semimajor axis of the orbit of the Moon increases, and the Moon drifts away about 3 cm per year.

7.5 Eclipses and Occultations

An *eclipse* is an event in which a body goes through the shadow of another body. The most frequently observed eclipses are the lunar eclipses and the eclipses of the large satellites of Jupiter. An *occultation* takes place when an occulting body goes in front of another object; typical examples are stellar occultations caused by the Moon. Generally, occultations can be seen only in a narrow strip; an eclipse is visible wherever the body is above the horizon.

Solar and lunar eclipses are the most spectacular events in the sky. A *solar eclipse* occurs when the Moon is between the Earth and the Sun (Fig. 7.7). (According to the definition, a solar eclipse is not an eclipse but an occultation!) If the whole disk of the Sun is behind the Moon, the eclipse is *total* (Fig. 7.8); otherwise, it is *partial*. If the Moon is close to its apogee, the apparent diameter of the Moon is smaller than that of the Sun, and the eclipse is *annular*.

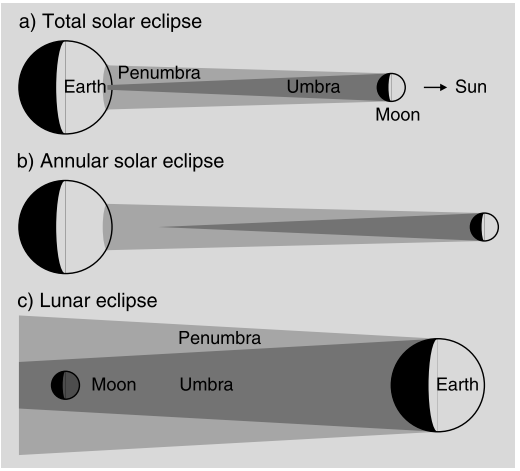


Fig. 7.7 (**a**) A total solar eclipse can be seen only inside a narrow strip; outside the zone of totality the eclipse is partial. (**b**) An eclipse is annular if the Moon is at apogee from where the shadow of the Moon does not reach the Earth. (**c**) A lunar eclipse is visible everywhere where the Moon is above the horizon

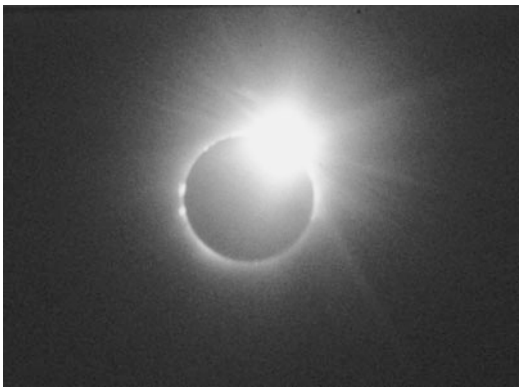


Fig. 7.8 The total eclipse of the Sun occurred in 1990 over Finland. (Photo Matti Martikainen)

If the orbital plane of the Moon coincided with the plane of the ecliptic, one solar and one lunar eclipse would occur every synodic month. However, the plane is tilted about 5°; therefore, at full moon, the Moon must be close to the nodes for an eclipse to occur. The angular distance of the Moon from the node must be smaller than 4.6°

for a total lunar eclipse, and 10.3° for a total solar eclipse.

Two to seven eclipses occur annually. Usually eclipses take place in a set of 1–3 eclipses, separated by an interval of 173 days. In one set there can be just one solar eclipse or a succession of solar, lunar and another solar eclipse. In one year, eclipses belonging to 2 or 3 such sets can take place.

The Sun and the (ascending or descending) node of the lunar orbit are in the same direction once every 346.62 days. Nineteen such periods (= 6585.78 days = 18 years 11 days) are very close to the length of 223 synodic months. This means that the Sun–Moon configuration and the eclipses are repeated in the same order after this period, called for historical reasons the *Saros* period; it was already known to the ancient Babylonians.

A solar eclipse is *total* if the whole solar disk is hidden behind the Moon. In a *partial* eclipse only a part of the Sun is covered. If the Moon is close to its apogee its apparent diameter is smaller than the diameter of the Sun, and the eclipse is *annular*.

During a solar eclipse the shadow of the Moon on Earth's surface is always less than 270 km wide. The shadow moves at least 34 km/min; thus the maximum duration of a total eclipse is $7\frac{1}{2}$ minutes. On both sides of the track of the total eclipse a partial eclipse can be seen.

Also lunar eclipses can be of different kinds. A lunar eclipse is *total* if the Moon is entirely inside the umbral shadow of the Earth; otherwise the eclipse is *partial*. If the Moon does not hit the umbral shadow but receives some sunlight, the eclipse is difficult to see with the unaided eye because the lunar magnitude remains almost unchanged.

A lunar eclipse can be seen on the whole hemisphere where the Moon is above the horizon. The maximum duration of a lunar eclipse is 3.8 h, and the duration of the total phase is always shorter than 1.7 h. During the total phase the Moon is coloured deep red because some red light is refracted through the Earth's atmosphere.

The Moon moves eastwards, and stars are occulted by the dark edge of the Moon during the first quarter. Therefore occultation is easier to observe, and photometric measurements are possible Observations of the *stellar occultations* caused by the Moon formerly served as an accurate method for determining the lunar orbit. Because the Moon has no atmosphere, a star disappears abruptly in less than 1/50 s. If a fast photometer is used for recording the event, the typical diffraction pattern can be seen. The shape of the diffraction was used to determine angular diameters of stars ans separations of binary stars. In the first decades of radio astronomy the occultations of some radio sources were used for determining their exact positions.

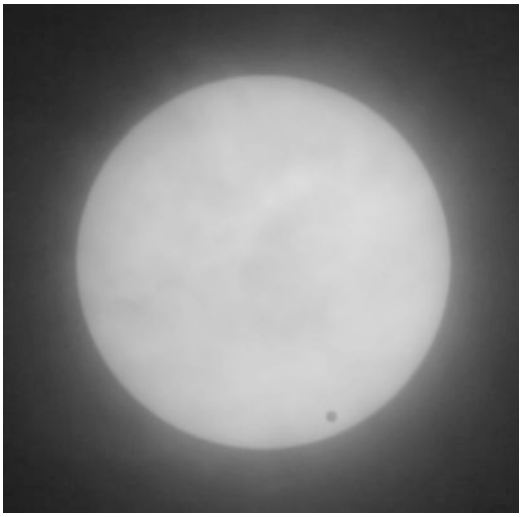


Fig. 7.9 Transit of Venus was seen in June 6, 2004. Venus is seen as a small dark disc in front of the Sun. (Photo H. Karttunen)

There are some bright stars and planets inside the 11° wide zone where the Moon moves, but the occultation of a bright, naked-eye object is quite rare.

Occultations are also caused by planets and asteroids. Accurate predictions are complicated because such an event is visible only in a very narrow path. The Uranian rings were found during an occultation in 1977, and the shapes of some asteroids have been studied during some favourable events, timed exactly by several observers located along the predicted path.

A *transit* is an event in which Mercury or Venus moves across the Solar disk as seen from the Earth (Fig. 7.9). A transit can occur only when

the planet is close to its orbital node at the time of inferior conjunction. Transits of Mercury occur about 13 times per century; transits of Venus only twice. The next transits of Mercury are: Nov 11, 2019; Nov 13, 2032 and Nov 7, 2039. The next transits of Venus are: Dec 11, 2117; Dec 8, 2125 and Jun 11, 2247. In the 18th century the two transits of Venus (1761 and 1769) were used for determining the value of the astronomical unit.

7.6 The Structure and Surfaces of Planets

Since the 1960's a vast amount of data have been collected using spacecraft, either during a flyby, orbiting a body, or directly landing on the surface. This gives a great advantage compared to other astronomical observations. We may even speak of revolution: the solar system bodies have turned from astronomical objects to geophysical ones. Many methods traditionally used in various sibling branches of geophysics can now be applied to planetary studies.

The perturbations in the orbit of a satellite or spacecraft can be used in studying the internal structure of a planet. Any deviation from spherical symmetry is visible in the external gravitational field. The shape and irregularities of the gravitation field generated by a planet reflect its shape, internal structure and mass distribution. Also the surface gives certain indications on internal structure and processes.

The IAU planet definition states that planets are bodies in *hydrostatic equilibrium*. Gravity of a body will pull its material inwards, but the body resist the pull if the strength of the material is greater than the pressure exerted by the overlying layers. If the diameter is larger than about 800–1000 km, gravity is able to deform rocky bodies into spherical shape. Smaller bodies than this have irregular shapes. On the other hand, e.g. icy moons of Saturn are spherical because ice is more easily deformed than rock.

Hydrostatic equilibrium means that the surface of the body approximately follows an equipotential surface of gravity. This is true e.g. on the Earth, where the sea surface very closely follows the equipotential surface called the *geoid*. Due to internal strength of rocks, continents can deviate from the geoid surface by a few kilometers but compared to the diameter of the Earth the surface topography is negligible.

A rotating planet is always *flattened*. The amount of flattening depends on the rotation rate and the strength of the material; a liquid drop is more easily deformed than a rock. The shape of a rotating body in hydrostatic equilibrium can be derived from the equations of motion. If the rotation rate is moderate, the equilibrium shape of a liquid body is an ellipsoid of revolution, rotating around its shortest axis.

If R_e and R_p are the equatorial and polar radii, respectively, the shape of the planet can be expressed as

$$\frac{x^2}{R_\mathrm{e}^2} + \frac{y^2}{R_\mathrm{e}^2} + \frac{z^2}{R_\mathrm{p}^2} = 1. \tag{7.2}$$

The *dynamical flattening*, denoted by f is defined as

$$f = \frac{R_\mathrm{e} - R_\mathrm{p}}{R_\mathrm{e}}. \tag{7.3}$$

Because $R_\mathrm{e} > R_\mathrm{p}$, the flattening f is always positive.

The giant planets are in practise close to hydrostatic equilibrium, and their shape is determined by the rotation. The rotation period of Saturn is only 10.5 h, and its dynamical flattening is $1/10$ which is easily visible.

Asteroids and other minor bodies are so small that they are not flattened by rotation. However, there is an upper limit for a rotation rate of an asteroid before it breaks apart due to centrifugal forces. If we assume that the body is held together only by gravity, we can approximate the maximum rotation rate by setting the centrifugal force equal to the gravitational force:

$$\frac{GMm}{R^2} = \frac{mv^2}{R}, \tag{7.4}$$

where m is a small test mass on the surface at a distance of R from the center of the body. Substituting the rotation period P,

$$P = \frac{2\pi R}{v},$$

we get

$$\frac{GM}{R^2} = \frac{4\pi^2 R}{P^2},$$

or

$$P = 2\pi\sqrt{\frac{R^3}{GM}} = 2\pi\sqrt{\frac{3}{4\pi G\rho}} = \sqrt{\frac{3\pi}{G\rho}}. \quad (7.5)$$

If we substitute the density ρ with the mean density of terrestrial rocks, i.e. 2700 kg m^{-3}, we get for the minimum rotation period $P \approx 2$ hours.

The structure of the terrestrial planets (Fig. 7.10) can also be studied with *seismic waves*. The waves formed in an earthquake are reflected and refracted inside a planet like any other wave at the boundary of two different layers. The waves are longitudinal or transversal (*P* and *S waves*, respectively). Both can propagate in solid materials such as rock. However, only the longitudinal wave can penetrate liquids. One can determine whether a part of the interior material is in the liquid state and where the boundaries of the layers are by studying the recordings of seismometers placed on the surface of a planet. Naturally the Earth is the best-known body, but quakes of the Moon, Venus, and Mars have also been observed.

The interior temperatures of the planets are considerably larger than the surface temperatures. For example, the temperature in the Earth's core is about 4500–5000 K, and in the core of Jupiter about 30,000 K.

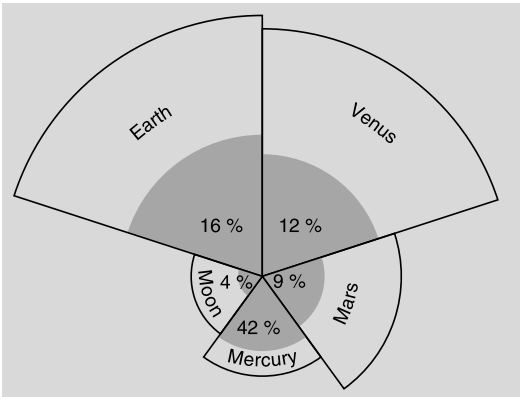


Fig. 7.10 Internal structure and relative sizes of the terrestrial planets. The percentage shows the volume of the core relative to the total volume of the planet. In the case of the Earth, the percentage includes both the outer and the inner core

A part of that heat is the remnant of the released potential energy from the gravitational contraction during the formation of planets. Decay of radioactive isotopes also releases heat. Soon after the formation of planets intense meteorite bombardment was an important source of heat. Together with heat from short-lived radioactive isotopes this caused *melting of terrestrial planets*. The planets were *differentiated*: the originally relatively homogeneous material became segregated into layers of different chemical composition. The heaviest elements sank into centre thus forming a dense core.

The terrestrial planets have an *iron-nickel core*. Mercury has the relatively largest core; Mars the smallest. The density of the core is around 10,000 kg m^{-3}. The Fe–Ni core is surrounded by a *mantle*, composed of *silicates* (compounds of silicon). The density of the outermost layers is about 3000 kg m^{-3}. The mean density of the terrestrial planets is 3500–5500 kg m^{-3}. A thin *crust* is the outermost layer. The crust and the upper mantle form a solid *lithosphere*, and the partly molten layer below that is the *asthenosphere*. Mean densities of planets are 3500–5500 kg m^{-3}, but the density of the surface material is less than 300 kg m^{-3}.

Rock is not a good heat conductor, but in terrestrial planets conduction is the only important method to transfer from the interior to the surface. Heat transfer by convection depends on the viscosity material and the temperature gradient. In the mantle of the Earth slow vertical convective flows occur below a few hundred kilometres. These flows drive e.g. the motion of the tectonic plates. Continental drift gives rise, for example, to mountain formation. The Earth is the only planet where *plate tectonics* is active today. On other terrestrial planets the process has either ceased long ago or has never occurred.

The giant planets (Fig. 7.11) do not have a similar isolating outer layer, and convection can transfer heat all the way to the surface. Thus the giant planets radiate more energy than they receive from the Sun. Saturn radiates about 2.8 times the heat it gets from the Sun, more than

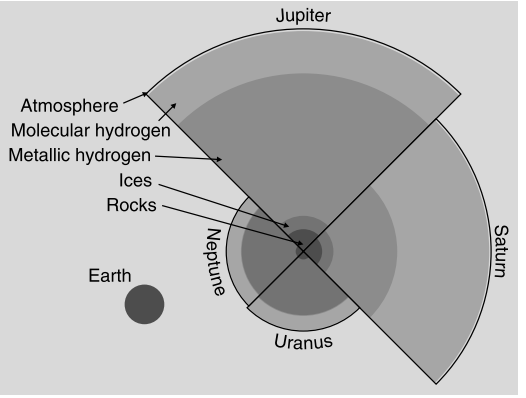


Fig. 7.11 Internal structure and relative sizes of the giant planets. Differences in size and distance from the Sun cause differences in the chemical composition and internal structure. Due to smaller size, Uranus and Neptune do not have any layer of metallic hydrogen. The Earth is shown in scale

any other planet. This heat is suspected to originate from the separation of hydrogen and helium, when the heavier helium is gradually sinking toward the centre of the planet. In Jupiter the heat is mainly remnant heat from the time the planet was born.

The mean densities of the giant planets are quite low; the density of Saturn, for example, is only 700 kg m^{-3}. (If Saturn were put in a gigantic bathtub, it would float on the water!) Most of the volume of a giant planet is a mixture of hydrogen and helium. In the centre, there is possibly a silicate core, the mass of which is a few Earth masses. The core is surrounded by a layer of *metallic hydrogen*. Due to the extreme pressure, hydrogen is not in its normal molecular form H_2, but dissociated into atoms. In this state, hydrogen is electrically conducting. The magnetic fields of the giant planets may originate in the layer of metallic hydrogen.

Closer to the surface, the pressure is lower and hydrogen is in molecular form. The relative thickness of the layers of metallic and molecular hydrogen vary from planet to planet. Uranus and Neptune may not have any layer of metallic hydrogen because their internal pressure is too low for dissociation of the hydrogen. Instead, a layer of "ices" surround the core. This is a layer of a water-dominant mixture of water, methane and ammonia. Under the high pressure and temperature the mixture is partly dissolved into its components and it behaves more like a molten salt and it is also electrically conductive like the metallic hydrogen.

On top of everything is a gaseous atmosphere, only a few hundred kilometres thick. The clouds at the top of the atmosphere form the visible "surface" of the giant planets.

Planetary surfaces are modified by several geological processes. These include *continental drift*, *volcanism*, *meteorite impacts* and *climate*. The Earth is an example of a body whose surface has been renewed many times during past aeons. The age of the surface depends on the processes and thus implies the geological evolutionary history of the planet.

Volcanism is a minor factor on the Earth (at least now), but the surface of the Jovian moon Io is changing rapidly due to violent volcanic eruptions (Fig. 7.12). Volcanism on Io is caused by frictional heating by tides. Volcanoes have been observed also on Mars and Venus, but not on the Moon.

Lunar craters are meteorite impact craters, common on almost every body with a solid surface. Meteorites are bombarding the planets continuously, but the rate has been diminishing since the beginnings of the solar system. The number of impact craters reflects the age of the surface (Figs. 7.13 and 7.14).

The Jovian moon Callisto is an example of a body with an ancient surface which is not fully inactive. Lack of small craters indicates some resurfacing process filling and degrading the minor surface features. The Earth is an example of a body, whose atmosphere both protects the surface and destroys the traces of impacts. All smaller meteorites are burned to ashes in the atmosphere (one need only note the number of shooting stars), and some larger bodies are bounced back to outer space. The traces on the surface are destroyed very quickly by erosion in less than a few million years. Venus is an even more extreme case where all small craters are missing due to a thick protective atmosphere.

Climate has the greatest influence on the Earth and Venus. Both planets have a thick atmosphere.

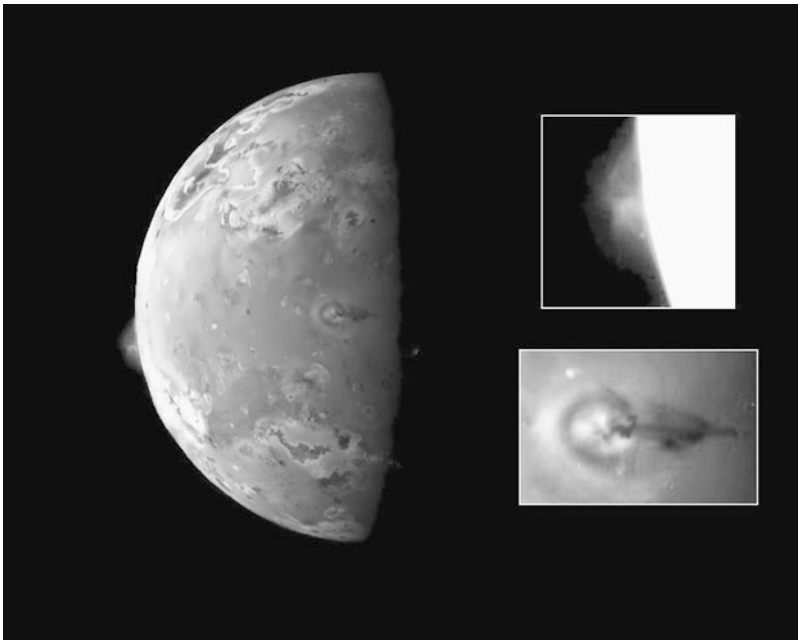


Fig. 7.12 An example of resurfacing. Two volcanic plumes on Jupiter's moon Io observed by Galileo spacecraft in 1997. One plume was captured on the bright limb or edge of the moon (*inset at upper right*), erupting over a caldera named Pillan Patera. The plume is 140 kilometers high. The second plume, seen near the terminator, is called Prometheus. The shadow of the 75 km high airborne plume can be seen extending to the right of the eruption vent. (NASA/JPL)

On Mars, powerful dust storms deform the landscape, too, often covering the planet with yellowish dust clouds.

7.7 Atmospheres and Magnetospheres

All major planets have an atmosphere; the atmosphere of Mercury, however, is extremely thin. The giant planets are surrounded by a very thick gaseous layer that can be regarded as an atmosphere. Saturn's moon Titan has a relatively thick atmosphere consisting of methane. Also the dwarf planet Pluto has a thin atmosphere consisting mainly of methane.

The composition, thickness, density and structure of the atmosphere vary from planet to planet, but some common features can be found (Figs. 7.15 and 7.16). Let us first study the dependence of the temperature T, pressure P, and density ρ on the height h. Let us consider a cylinder with a length $\mathrm{d}h$. The change in the pressure $\mathrm{d}P$ from the height h to $h + \mathrm{d}h$ is proportional to the mass of the gas in the cylinder:

$$\mathrm{d}P = -g\rho\,\mathrm{d}h, \tag{7.6}$$

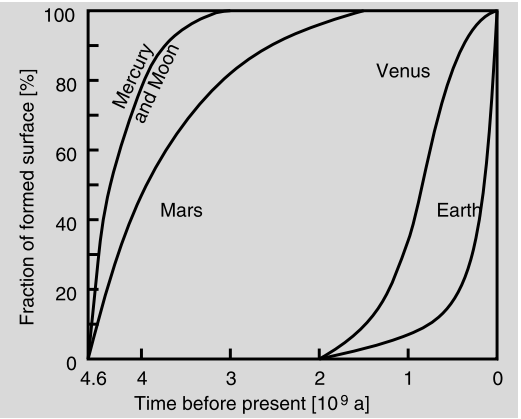


Fig. 7.13 Ages of the surfaces of Mercury, the Earth, the Moon and Mars. *The curve* represents the fraction of the surface which existed at a certain time. Most of the surface of the Moon, Mercury and Mars are more than 3500 million years old, whereas the surface of the Earth is mostly younger than 200 million years

where g is the acceleration of gravity. Equation (7.6) is the *equation of hydrostatic equilibrium*. (It is discussed in detail in Chap. 11.)

As a first approximation, we may assume that g does not depend on height. In the case of the Earth, the error is only about 3 % if g is considered constant from the surface to a height of 100 km.

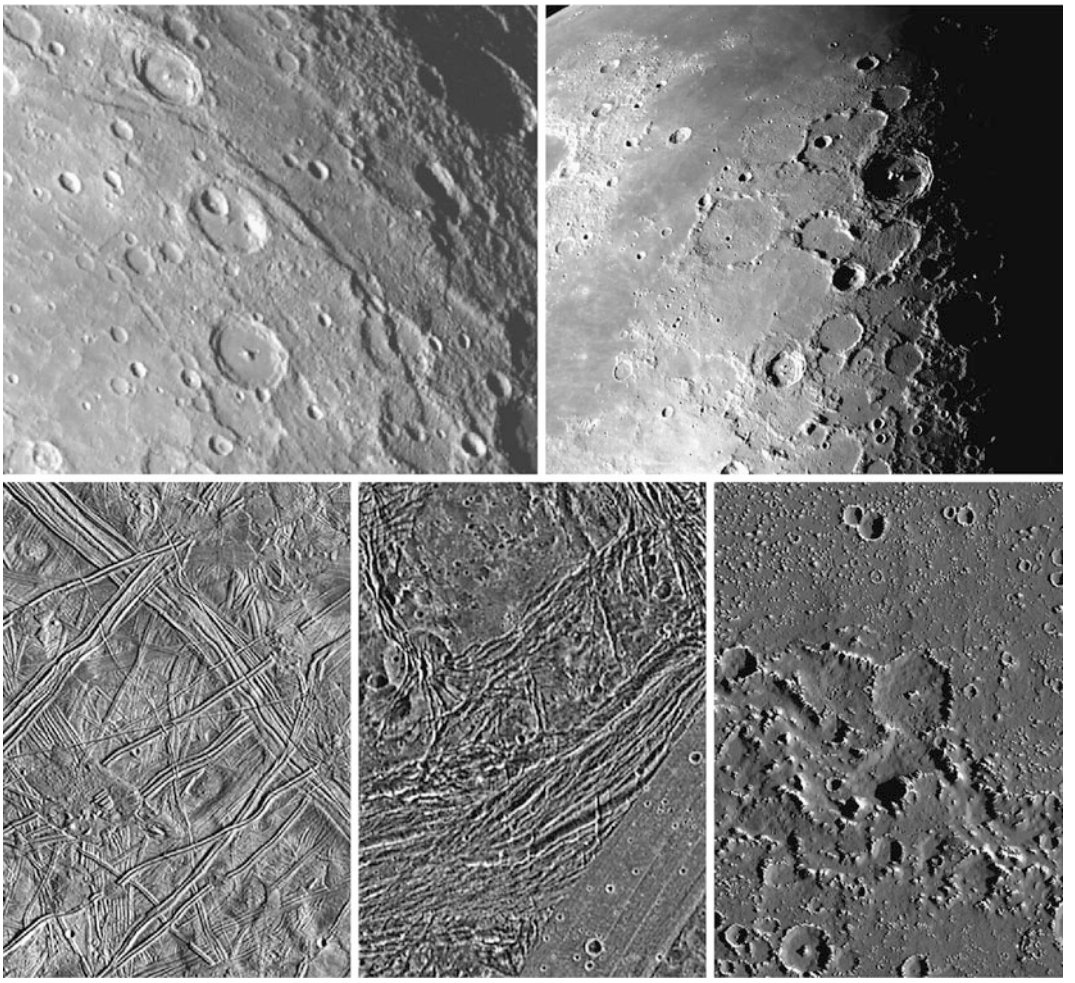


Fig. 7.14 The number of meteorite impact craters is an indicator of the age of the surface and the shapes of the craters give information on the strength of the material. *The upper row* shows Mercury (*left*) and the Moon, and *the second row*, the Jovian moons Europa (*left*), Ganymede (*centre*) and Callisto. The pictures of the Jovian moons were taken by the Galileo orbiter with a resolution of 150 metres/pixel. Europa has only a few craters, there are areas of different ages on the surface Ganymede and the surface of Callisto is the oldest. Note the grooves and ridges that indicate different geological processes. IN the bottom there are two volcanic plumes on Jupiter's moon Io observed by Galileo spacecraft in 1997. One plume was captured on the bright limb or edge of the moon (*inset at upper right*), erupting over a caldera named Pillan Patera. The plume is 140 kilometers high. The second plume, seen near the terminator, is called Prometheus. The shadow of the 75 km high airborne plume can be seen extending to the right of the eruption vent. (NASA/JPL and DLR)

The equation of state of the ideal gas

$$PV = NkT \tag{7.7}$$

gives the expression for the pressure P

$$P = \frac{\rho kT}{\mu}, \tag{7.8}$$

where N is the number of atoms or molecules, k is the Boltzmann constant, μ is the mass of one atom or molecule and

$$\rho = \frac{\mu N}{V}.$$

By using the equation of hydrostatic equilibrium (7.6) and the equation of state (7.8), we obtain

$$\frac{\mathrm{d}P}{P} = -g\frac{\mu}{kT}\,\mathrm{d}h.$$

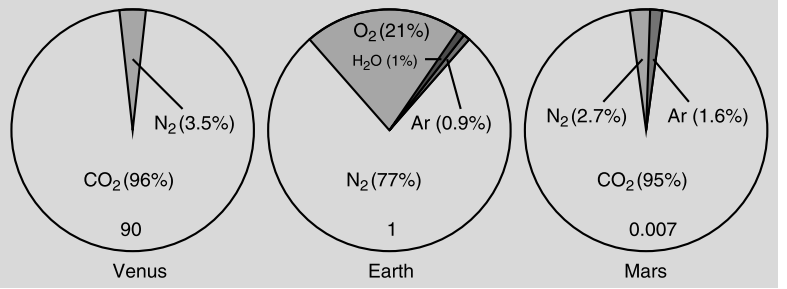


Fig. 7.15 Relative abundances of the most abundant gases in the atmospheres of Venus, Earth, and Mars. The number at the bottom of each circle denotes the surface pressure in atms

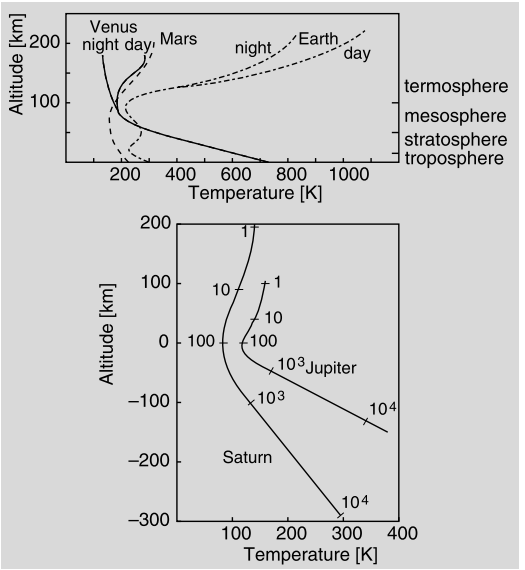


Fig. 7.16 (**a**) Temperature as a function of height in the atmospheres of Venus, Earth, and Mars. (**b**) Temperature profiles of the atmospheres of Jupiter and Saturn. The zero height is chosen to be the point where the pressure is 100 mbar. Numbers along the curves are pressures in millibars

Integration yields P as a function of height:

$$P = P_0 \exp\left(-\int_0^h \frac{\mu g}{kT}\,\mathrm{d}h\right) = P_0 \exp\left(-\int_0^h \frac{\mathrm{d}h}{H}\right). \tag{7.9}$$

The variable H, which has the dimension of length, is called the *scale height*:

$$H = \frac{kT}{\mu g}. \tag{7.10}$$

The scale height defines the height at which the pressure has decreased by a factor e. H is a function of height, but here we may assume that it is constant. With this approximation, we obtain

$$-\frac{h}{H} = \ln\frac{P}{P_0}$$

or, using (7.8),

$$\frac{\rho T(h)}{\rho_0 T_0} = \mathrm{e}^{-h/H}. \tag{7.11}$$

The scale height is an important parameter in many formulas describing the structure of the atmosphere (Table 7.1). For example, if the change of the pressure or the density is known as a function of height, the mean molecular weight of the atmosphere can be computed. The scale height of the Jovian atmosphere was determined in 1952 when Jupiter occulted a star. With these observations, the scale height was calculated to be 8 km, and the mean molecular weight 3–5 amu (atomic mass unit, 1/12 of the mass of ^{12}C). Thus the main components are hydrogen and helium, a result later confirmed by spacecraft data.

In terrestrial observations, infrared data are limited by water vapour and carbon dioxide. The scale height of CO_2 is 5 km, which means that the partial pressure is already halved at a height of 3.5 km. Thus infrared observations can be made on top of high mountains (like Mauna Kea in Hawaii). The scale height of water vapour is 13 km, but the relative humidity and hence the actual water content is very site- and time-dependent.

The scale height and the temperature of the atmosphere define the permanence of the atmosphere. If the speed of a molecule is greater than

Table 7.1 Scale heights of some gases in the atmospheres of Venus, Earth, and Mars

Gas	Molecular weight [amu]	Earth H [km]	Venus H [km]	Mars H [km]
H_2	2	120	360	290
O_2	32	7	23	18
H_2O	18	13	40	32
CO_2	44	5	16	13
N_2	28	8	26	20
Temperature [K]		275	750	260
Acceleration of gravity [m/s^2]		9.81	8.61	3.77

the escape velocity, the molecule will escape into space. The whole atmosphere could disappear in a relatively short time.

According to the kinetic gas theory, the mean velocity $\overline{v}$ of a molecule depends both on the kinetic temperature T_k of the gas and the mass m of the molecule:

$$\overline{v} = \sqrt{\frac{3kT_k}{m}}. \tag{7.12}$$

If the mass of a planet is M and its radius R, the escape velocity is

$$v_e = \sqrt{\frac{2GM}{R}}. \tag{7.13}$$

Even if the mean velocity is smaller than the escape velocity, the atmosphere can evaporate into space if there is enough time, since some molecules will always have velocities exceeding v_e. Assuming a velocity distribution, one can calculate the probability for $v > v_e$ Hence it is possible to estimate what fraction of the atmosphere will disappear in, say, 10^9 years. As a rule of thumb, it can be said that at least half of the atmosphere will remain over 1000 million years if the mean velocity $\overline{v} < 0.2v_e$.

Giant planets move far from the Sun; thus the surface temperature is low. Also, the gravitation is strong. Thus it is understandable that e.g. Jupiter has been able to retain more hydrogen than the Earth.

The probability that a molecule close to the surface will escape is insignificantly small. The free mean path of a molecule is very small when the gas density is high (Fig. 7.17). Thus the escaping molecule is most probably leaving from the uppermost layers. The *critical layer* is defined as a height at which a molecule, moving upward, has a probability $1/e$ of hitting another molecule. The part of the atmosphere above the critical layer is called the *exosphere*. The exosphere of the Earth begins at a height of 500 km, where the kinetic temperature of the gas is 1500–2000 K and the pressure is lower than in the best terrestrial vacuums.

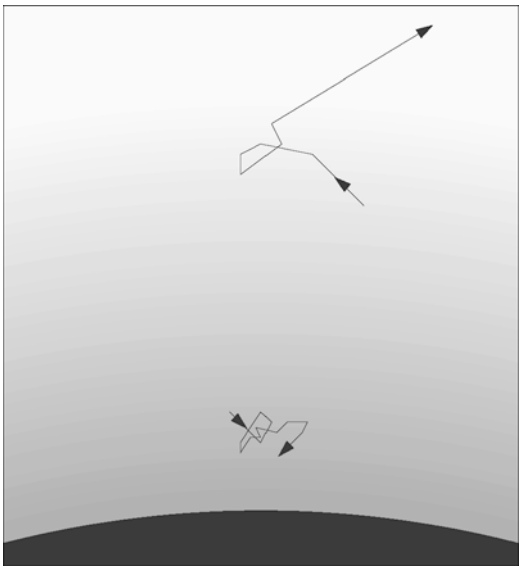


Fig. 7.17 Close to the surface, the mean free path of a molecule is smaller than higher in the atmosphere where the gas density is smaller. The escaping molecules originate close to the critical layer

The *magnetosphere* is the "outer boundary" of a planet. Size and shape depend on the strength of the magnetic field of the planet and on the solar wind. The *solar wind* is a flux of charged particles, mostly electrons and protons, outflow-

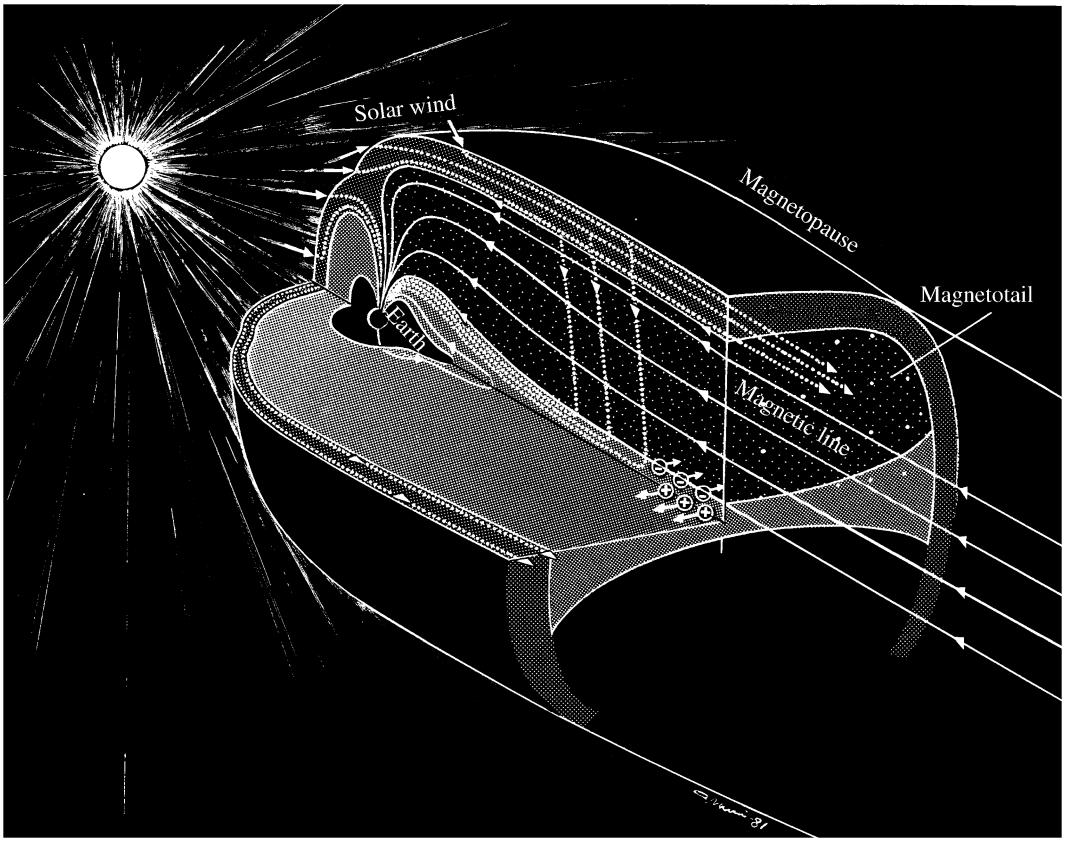


Fig. 7.18 Structure of the magnetosphere of the Earth. (A. Nurmi/Tiede 2000)

ing from the Sun. The speed of the wind at the distance of the Earth is about 500 km/s and the density 5–10 particles/cm^3 but both values can change considerably depending on the solar activity.

On the solar side there is a *bow shock* (Fig. 7.18), typically at a distance of a few tens of planetary radii (Table 7.2). At the bow shock, particles of the solar wind first hit the magnetosphere. The magnetosphere is limited by the *magnetopause*, flattened on the solar side and extended to a long tail on the opposite side. Charged particles inside the magnetopause are captured by the magnetic field and some particles are accelerated to great velocities (Fig. 7.19). If the velocities are interpreted according to the kinetic gas theory, these velocities even correspond to millions of kelvins. However, the density, and thus the total energy, is very small. The "hottest" places found are around Jupiter and Saturn, where the particle velocities correspond even temperatures of 400 million kelvins.

The region of space containing trapped charged particles, the radiation belts around the Earth, are named *van Allen's belts*. These radiation zones were discovered by the first US satellite, Explorer 1, in 1958.

The number of charged particles increases after strong solar bursts. Some of these particles "leak" to the atmosphere, resulting in *auroras*. Similar effects have also been detected in Jupiter, Saturn and Uranus.

The solar magnetic field arises from the turbulent motions of the electrically conductive matter. The energy driving the convection in the layer is coming from the nuclear fusion in the core. This, however, cannot explain planetary magnetism. Neither can the remanent primordial magnetic

Table 7.2 Planetary magnetic fields

	Dipole moment (Earth = 1)	Field strength (gauss)[a]	Polarity[b]	Angle[c]	Magneto-pause[d]
Mercury	0.0007	0.003	⇑	14°	1.5
Venus	<0.0004	<0.00003	–	–	–
Earth	1.0	0.305	⇑	11°	10
Mars	<0.0002	<0.0003	–	–	–
Jupiter	20,000.	4.28	⇓	10°	80
Saturn	600.	0.22	⇓	<1°	20
Uranus	50.	0.23	⇓	59°	20
Neptune	25.	0.14	⇓	47°	25

[a] At equator (1 gauss equals 10^{-4} T)

[b] ⇑ same as the Earth, ⇓ opposite

[c] Angle between magnetic and rotational axes

[d] Average magnetopause distance in the direction of the Sun in planetary radii

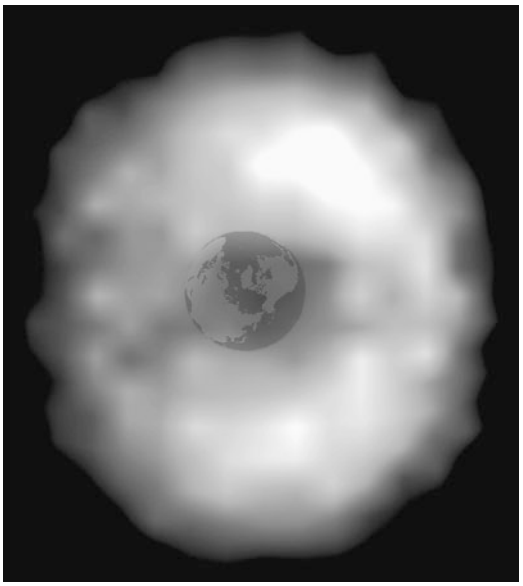


Fig. 7.19 A glow of hot plasma trapped inside the Earth's magnetosphere. The picture was taken by NASA's Imager for Magnetopause to Aurora Global Exploration (IMAGE) spacecraft on August 11, 2000 at 18:00 UT. The Sun is outside the picture area toward the top right corner. (NASA and the IMAGE science team)

field explain it because the internal temperature of planets is well above the *Curie point* (about 850 K for magnetite). If the temperature is above the Curie point, ferromagnetic materials will lose their remanent magnetism.

The planetary *dynamo* generating the magnetic field requires that the planet is rotating and has a convective layer of electrically conductive material. The temperature gradient across this layer must also be high enough to maintain convection. Terrestrial planets have a liquid Fe–Ni core, or a liquid layer in the core, Jupiter and Saturn have a layer of liquid metallic hydrogen and Uranus and Neptune have a mixture of water, ammonia and methane.

The strength of the magnetic field varies a lot from planet to planet. It can be characterised by the *dipole magnetic moment*. The magnetic moment of Jupiter is about 100 million times that of Mercury. The magnetic moment of the Earth is about 7.9×10^{25} gauss cm^3 that can be compared to the typical strong electromagnetic fields achieved in the laboratories, about 100,000 gauss cm^3. Inducing such a strong field requires currents that are of the order of 10^9 Amperes. When divided by the cube of planetary radii, one gets an estimate of the field strength on the equator.

The alignment of the magnetic field with respect to the rotation axis of a planet differs from planet to planet (Fig. 7.20). Saturn's magnetic field is close to the ideal case where rotational axis and magnetic axis coincide. Also the Earth and Jupiter show reasonably good point dipole field with a tilt of about 10°. However, fields of

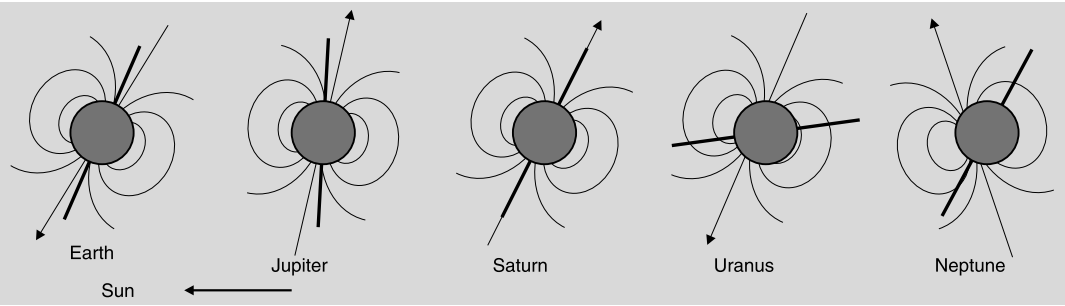


Fig. 7.20 Planetary magnetic fields

Uranus and Neptune are both offset from the centre of the planet and tilted by about 50° from the rotation axis. This may indicate a different mechanism for the dynamo.

The magnetic fields of Mercury and the Earth have an opposite *polarity* than the fields of other planets. It is known that the polarity of the Earth's magnetic field has reversed several times over geologic time scales, previously about 750,000 years ago. There are some indications that the reversal of the polarity is beginning now because the field strength is declining about one percent per decade, magnetic poles are moving more rapidly and the field asymmetry is increasing. The whole process will take several thousand years during which the Earth's surface is more open to the cosmic rays.

The Galileo mission also revealed that the Jovian moon *Ganymede* has a magnetic field. The field is weak and too small to have a magnetotail or trapped particles around the moon. Callisto, which is of the same size, does not show any magnetosphere. Neither does our Moon have any global magnetic field.

7.8 Albedos

The planets and all other bodies of the solar system only reflect the radiation of the Sun (we may neglect here the thermal and radio wave radiation and concentrate mainly on the visual wavelengths). The brightness of a body depends on its distance from the Sun and the Earth, and on the albedo of its surface. The term *albedo* defines the ability of a body to reflect light.

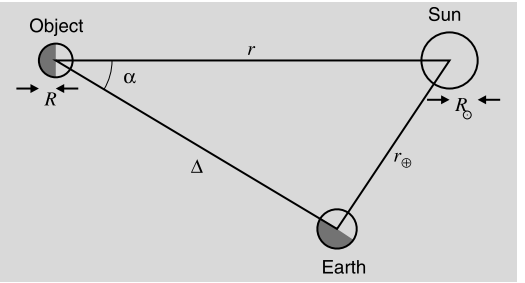


Fig. 7.21 Symbols used in the photometric formulas. The angle α is the phase angle

If the luminosity of the Sun is $L_\odot$, the flux density at the distance r is (Fig. 7.21)

$$F = \frac{L_\odot}{4\pi r^2}. \tag{7.14}$$

If the radius of the planet is R, the area of its cross section is πR^2, and the total flux incident on the surface of the planet is

$$L_{\text{in}} = \pi R^2 \frac{L_\odot}{4\pi r^2} = \frac{L_\odot R^2}{4r^2}. \tag{7.15}$$

Only a part of the incident flux is reflected back. The other part is absorbed and converted into heat which is then emitted as a thermal emission from the planet. The *Bond albedo* A (or spherical albedo) is defined as the ratio of the emergent flux to the incident flux ($0 \le A \le 1$). The flux reflected by the planet is thus

$$L_{\text{out}} = AL_{\text{in}} = \frac{AL_\odot R^2}{4r^2}. \tag{7.16}$$

The planet is observed at a distance Δ. If radiation is reflected isotropically, the observed flux

density should be

$$F = \frac{L_{\text{out}}}{4\pi\,\Delta^2}. \tag{7.17}$$

In reality, however, radiation is reflected anisotropically. If we assume that the reflecting object is a homogeneous sphere, the distribution of the reflected radiation depends on the *phase angle* α only. Thus we can express the flux density observed at a distance Δ as

$$F = C\Phi(\alpha)\frac{L_{\text{out}}}{4\pi\,\Delta^2}. \tag{7.18}$$

The function Φ giving the phase angle dependence is called the *phase function*. It is normalised so that $\Phi(\alpha = 0^\circ) = 1$.

Since all the radiation reflected from the planet is found somewhere on the surface of the sphere, we must have

$$\int_S C\Phi(\alpha)\frac{L_{\text{out}}}{4\pi\,\Delta^2}\,\mathrm{d}S = L_{\text{out}} \tag{7.19}$$

or

$$\frac{C}{4\pi\,\Delta^2}\int_S \Phi(\alpha)\,\mathrm{d}S = 1, \tag{7.20}$$

where the integration is extended over the surface of the sphere of radius Δ. The surface element of such a sphere is $\mathrm{d}S = \Delta^2\,\mathrm{d}\alpha\sin\alpha\,\mathrm{d}\phi$, and we have

$$\begin{aligned}\int_S \Phi(\alpha)\,\mathrm{d}S &= \Delta^2\int_{\alpha=0}^{\pi}\int_{\phi=0}^{2\pi}\Phi(\alpha)\sin\alpha\,\mathrm{d}\alpha\,\mathrm{d}\phi \\ &= \Delta^2 2\pi\int_0^{\pi}\Phi(\alpha)\sin\alpha\,\mathrm{d}\alpha.\end{aligned} \tag{7.21}$$

The normalisation constant C is

$$C = \frac{4\pi\,\Delta^2}{\int_S \Phi(\alpha)\,\mathrm{d}S} = \frac{2}{\int_0^{\pi}\Phi(\alpha)\sin\alpha\,\mathrm{d}\alpha}. \tag{7.22}$$

The quantity

$$q = 2\int_0^{\pi}\Phi(\alpha)\sin\alpha\,\mathrm{d}\alpha \tag{7.23}$$

is the *phase integral*. In terms of the phase integral the normalisation constant is

$$C = \frac{4}{q}. \tag{7.24}$$

Remembering that $L_{\text{out}} = AL_{\text{in}}$, (7.18) can be written in the form

$$F = \frac{CA}{4\pi}\Phi(\alpha)\frac{1}{\Delta^2}L_{\text{in}}. \tag{7.25}$$

The first factor is intrinsic for each object, the second gives the phase angle dependence, the third the distance dependence and the fourth, the incident radiation power. The first factor is often denoted by

$$\Gamma = \frac{CA}{4\pi}. \tag{7.26}$$

When we substitute here the expression of C (7.24), and solve for the Bond albedo, we get

$$A = \frac{4\pi\Gamma}{C} = \pi\Gamma\frac{4}{C} = \pi\Gamma q = pq. \tag{7.27}$$

Here $p = \pi\Gamma$ is called the *geometric albedo* and q is the previously introduced phase integral. These quantities are related by

$$A = pq. \tag{7.28}$$

The geometric albedo seems to have appeared as an arbitrary factor with no obvious physical interpretation. We'll now try to explain this quantity using a *Lambertian surface*. A Lambertian surface is defined as an absolutely white, diffuse surface which reflects all radiation, i.e. its Bond albedo is $A = 1$. Moreover, its surface brightness is the same for all viewing directions, which means that the phase function is

$$\Phi(\alpha) = \begin{cases}\cos\alpha, & \text{if } 0 \le \alpha \le \pi/2,\\ 0, & \text{otherwise.}\end{cases} \tag{7.29}$$

In reality, no such surface exists but there are some materials which behave almost like a Lambertian surface. A wall with a mat white finish is a good approximation; although it doesn't reflect all incident light, the distribution of the reflected light is about right, and its brightness looks the same from all directions.

For a Lambertian surface the constant C is

$$C = \frac{2}{\int_0^{\pi}\Phi(\alpha)\sin\alpha\,\mathrm{d}\alpha}$$

$$= \frac{2}{\int_0^{\pi/2} \cos\alpha \sin\alpha \, \mathrm{d}\alpha}$$

$$= \frac{2}{1/2} = 4. \tag{7.30}$$

Thus the geometric albedo of a Lambertian surface is

$$p = \pi\Gamma = \frac{CA}{4} = \frac{4 \times 1}{4} = 1. \tag{7.31}$$

At the phase angle zero $\Phi(\alpha = 0^\circ) = 1$ and the reflected flux density is

$$F = \frac{CA}{4\pi} \frac{1}{\Delta^2} L_{\mathrm{in}}.$$

If we replace the object with a Lambertian surface of the same size, we get

$$F_{\mathrm{L}} = \frac{4}{4\pi} \frac{1}{\Delta^2} L_{\mathrm{in}}.$$

The ratio of these flux densities is

$$\frac{F}{F_{\mathrm{L}}} = \frac{CA}{4} = \pi\Gamma = p. \tag{7.32}$$

Now we have found a physical interpretation for p: the geometric albedo is the ratio of the flux densities at phase angle $\alpha = 0^\circ$ reflected by a planet and a Lambertian surface of the same cross section.

The geometric albedo depends on the reflectance of the surface but also on the phase function Φ. Many rough surfaces reflect most of the incident radiation directly backward. In such a case the geometric albedo p is greater than in the case of an isotropically reflecting surface. On some surfaces $p > 1$; the most extreme case is a mirror, for which the specular reflection, $p = \infty$. The geometric albedo of solar system bodies vary between 0.03–1. The geometric albedo of the Moon is $p = 0.12$ and the greatest value, $p = 1.0$, has been measured for the Saturnian moon Enceladus.

It turns out that p can be derived from the observations, but the Bond albedo A can be determined only if the phase integral q is also known. That will be discussed in the next section.

7.9 Photometry, Polarimetry and Spectroscopy

Having defined the phase function and albedos we are ready to derive a formula for *planetary magnitudes*. The flux density of the reflected light is

$$F = \frac{CA}{4\pi} \Phi(\alpha) \frac{1}{\Delta^2} L_{\mathrm{in}}.$$

We now substitute the incident flux

$$L_{\mathrm{in}} = \frac{L_\odot R^2}{4r^2}$$

and the constant factor expressed in terms of the geometric albedo

$$\frac{CA}{4\pi} = \Gamma = \frac{p}{\pi}.$$

Thus we get

$$F = \frac{p}{\pi} \Phi(\alpha) \frac{1}{\Delta^2} \frac{L_\odot R^2}{4r^2}. \tag{7.33}$$

The observed solar flux density at a distance of $a = 1$ au from the Sun is

$$F_\odot = \frac{L_\odot}{4\pi a^2}. \tag{7.34}$$

The ratio of these is

$$\frac{F}{F_\odot} = \frac{p\Phi(\alpha) R^2 a^2}{\Delta^2 r^2}. \tag{7.35}$$

If the apparent solar magnitude at a distance of 1 au is $m_\odot$ and the apparent magnitude of the planet m we have

$$\begin{aligned} m - m_\odot &= -2.5 \lg \frac{F}{F_\odot} \\ &= -2.5 \lg \frac{p\Phi(\alpha) R^2 a^2}{\Delta^2 r^2} \\ &= -2.5 \lg \frac{pR^2}{a^2} \frac{a^4}{\Delta^2 r^2} \Phi(\alpha) \\ &= -2.5 \lg p \frac{R^2}{a^2} - 2.5 \lg \frac{a^4}{\Delta^2 r^2} \\ &\quad - 2.5 \lg \Phi(\alpha) \end{aligned}$$

$$= -2.5 \lg p \frac{R^2}{a^2} + 5 \lg \frac{\Delta r}{a^2} - 2.5 \lg \Phi(\alpha). \quad (7.36)$$

If we denote

$$V(1,0) \equiv m_\odot - 2.5 \lg p \frac{R^2}{a^2}, \quad (7.37)$$

then the magnitude of a planet can be expressed as

$$m = V(1,0) + 5 \lg \frac{r\Delta}{a^2} - 2.5 \lg \Phi(\alpha). \quad (7.38)$$

The first term $V(1,0)$ depends only on the size of the planet and its reflection properties. So it is a quantity intrinsic to the planet, and it is called the *absolute magnitude* (not to be confused with the absolute magnitude in stellar astronomy!). The second term contains the distance dependence and the third one the dependence on the phase angle.

If the phase angle is zero, and we set $r = \Delta = a$, (7.38) becomes simply $m = V(1,0)$. The absolute magnitude can be interpreted as the magnitude of a body if it is at a distance of 1 au from the Earth and the Sun at a phase angle $\alpha = 0°$. As will be immediately noticed, this is physically impossible because the observer would be in the very centre of the Sun. Thus $V(1,0)$ can never be observed.

By using (7.37) and (7.38) at $\alpha = 0°$, the geometric albedo can be solved for in terms values all obtainable from observations.

$$p = \left(\frac{r\Delta}{aR}\right)^2 10^{-0.4(m_0 - m_\odot)}, \quad (7.39)$$

where $m_0 = m(\alpha = 0°)$. As can easily be seen, p can be greater than unity but in the real world, it is normally well below that. Typical values for p are in the range 0.1–0.5.

The last term containing the phase angle dependence in (7.38) is the most problematic one. For many objects the phase function is not known very well. This means that from the observations, one can calculate only

$$V(1,\alpha) \equiv V(1,0) - 2.5 \lg \Phi(\alpha), \quad (7.40)$$

which is the *absolute magnitude at phase angle* α. $V(1,\alpha)$, plotted as a function of the phase angle, is called the *phase curve* (Fig. 7.22). The phase curve extrapolated to $\alpha = 0°$ gives $V(1,0)$. The shape of the phase curve is very different for objects with or without an atmosphere.

The Bond albedo can be determined only if the phase function Φ is known. Superior planets (and other bodies orbiting outside the orbit of the Earth) can be observed only in a limited phase angle range, and therefore Φ is poorly known, except for those bodies that have been observed by spacecraft. The situation is somewhat better for the inferior planets. Especially in popular texts the Bond albedo is given instead of p (naturally without mentioning the exact names!). A good excuse for this is the obvious physical meaning of the former, and also the fact that the Bond albedo is normalised to [0, 1].

Opposition Effect If an object has an atmosphere it reflects light more or less isotropically to all directions. The flux density of the reflected light is then proportional to the area of the visible illuminated surface (actually to the projection of this area on a plane perpendicular to the line of sight). Atmosphereless bodies reflect light more strongly to the direction of the incident light. Hence the brightness increases rapidly when the phase angle approaches zero. When the phase is larger than about 10°, the changes are smaller. This rapid brightening close to the opposition is called the *opposition effect*. An atmosphere destroys the opposition effect.

The full explanation is still in dispute. A qualitative (but only partial) explanation is that close to the opposition, no shadows are visible. When the phase angle increases, the shadows become visible and the brightness drops. The main reason, however, is the coherent backscatter due to the wave properties of the light.

Magnitudes of Asteroids The shape of the phase curve depends on the geometric albedo. It is possible to estimate the geometric albedo if the phase curve is known. This requires at least a few observations at different phase angles. Most critical is the range 0°–10°. A known phase curve

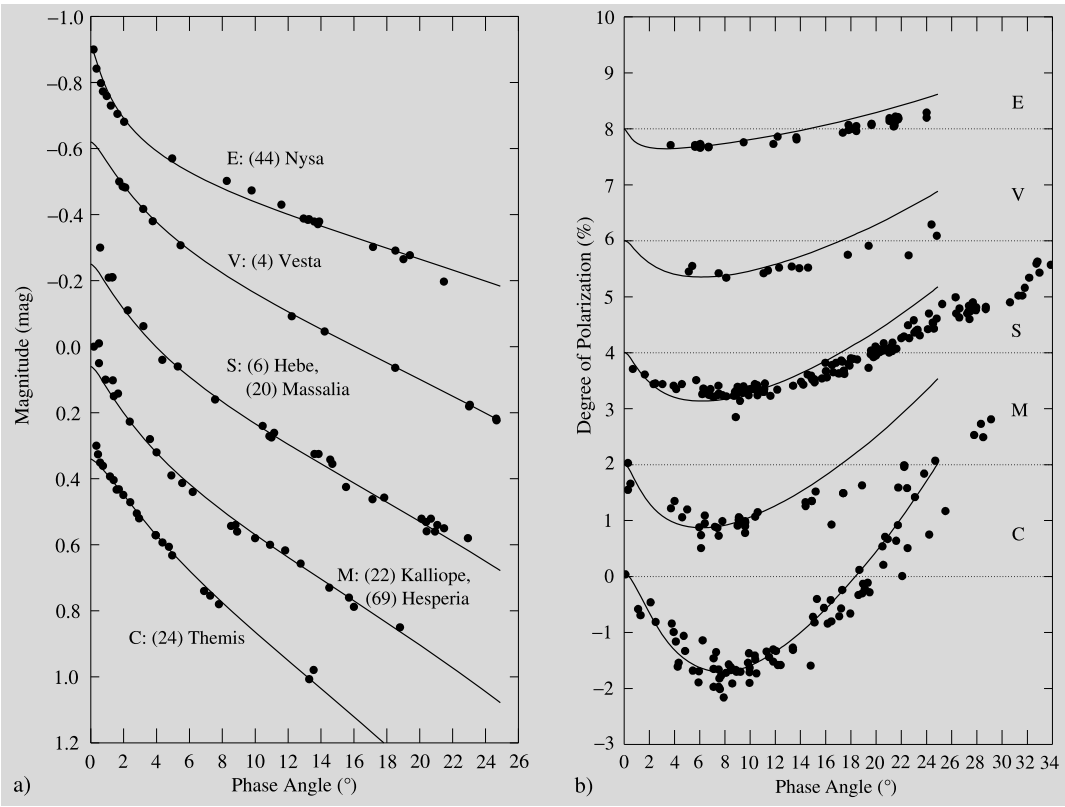


Fig. 7.22 The phase curves and polarisation of different types of asteroids. The asteroid characteristics are discussed in more detail in Sect. 8.11. (From Muinonen *et al.*, *Asteroid photometric and polarimetric phase effects*, in Bottke, Binzel, Cellino, Paolizhi (Eds.) *Asteroids III*, University of Arizona Press, Tucson)

can be used to determine the diameter of the body, e.g. the size of an asteroid. Apparent diameters of asteroids are so small that for ground based observations one has to use indirect methods, like polarimetric or radiometric (thermal radiation) observations (Fig. 7.22). Beginning from the 1990's, imaging made during spacecraft flybys and with the Hubble Space Telescope have given also direct measures of the diameter and shape of asteroids.

When the phase angle is greater than a few degrees, the magnitude of an asteroid depends almost linearly on the phase angle. Earlier this linear part was extrapolated to $\alpha = 0°$ to estimate the opposition magnitude of an asteroid. Due to the opposition effect the actual opposition magnitude can be considerably brighter.

In 1985 the IAU adopted the *HG system* for magnitudes of atmosphereless bodies. Formally, it was semi-empirical, although it was based on photometric theories by Lumme and Bowell. In the 2012 meeting this was replaced by a new HG_1G_2 system. Although the older HG system was useful in many cases, it was not satisfactory if the opposition effect was very small or restricted to very small phase angles.

In the new system the magnitude at phase angle α is

$$\begin{aligned} V(1,\alpha) &= -2.5\lg\big[a_1\Phi_1(\alpha) + a_2\Phi_2(\alpha) + a_3\Phi_3(\alpha)\big] \\ &= H - 2.5\lg\big[G_1\Phi_1(\alpha) + G_2\Phi_2(\alpha) \\ &\quad + (1 - G_1 - G_2)\Phi_3(\alpha)\big], \end{aligned} \tag{7.41}$$

where the values of the basis functions Φ_1, Φ_2 and Φ_3 are found by spline interpolations from the following tables:

a [°]	Φ_1
0.0	1.0
7.5	0.75
30.0	0.33486016
60.0	0.13410560
90.0	0.05110476
120.0	0.02146569
150.0	0.00363970
$\Phi_1'(7.5°) = -1.90986$	
$\Phi_1'(150°) = -0.09133$	

a [°]	Φ_2
0.0	1.0
7.5	0.925
30.0	0.62884169
60.0	0.31755495
90.0	0.12716367
120.0	0.02237390
150.0	0.00016506
$\Phi_2'(7.5°) = -0.57330$	
$\Phi_2'(150°) = -8.657 \times 10^{-8}$	

a [°]	Φ_3
0.0	1.0
0.3	0.83381185
1.0	0.57735424
2.0	0.42144772
4.0	0.23174230
8.0	0.10348178
12.0	0.06173347
20.0	0.01610701
30.0	0.0
$\Phi_3'(0°) = -0.10630$	
$\Phi_3'(30°) = 0$	

When the phase angle is zero all the functions have the value 1. If the phase angle is $\alpha \leq 7.5°$, Φ_1 ans Φ_2 are linear functions: $\Phi_1(\alpha) = 1 - \alpha/30°$, $\Phi_2(\alpha) = 1 - \alpha/100°$.

Fitting an expression in terms of the basis functions to the observed phase curve one gets the coefficients a_i, and then further

$$\begin{aligned} H &= -2.5\lg(a_1 + a_2 + a_3), \\ G_1 &= a_1/(a_1 + a_2 + a_3), \\ G_2 &= a_2/(a_1 + a_2 + a_3), \end{aligned} \tag{7.42}$$

When the phase angle is zero, we have

$$\begin{aligned} V(1,0) &= H - 2.5\lg[G_1 + G_2 + 1 - G_1 - G_2] \\ &= H, \end{aligned} \tag{7.43}$$

and hence H is just the absolute magnitude in opposition. The constants G_1 and G_2 describe the shape of the phase curve.

Asteroid data has earlier been published in the yearbook *Efemeridy malyh planet*. Currently the best source is the web pages of the Minor Planet Center: http://www.cfa.harvard.edu/iau/services/WebCSAccess.html.

Polarimetric Observations The light reflected by the bodies of the solar system is usually polarised, at least to some degree. The amount of polarisation depends on the reflecting material and also on the geometry: polarisation is a function of the phase angle. The *degree of polarisation* P is defined as

$$P = \frac{F_\perp - F_\parallel}{F_\perp + F_\parallel}, \tag{7.44}$$

where $F_\perp$ is the flux density of radiation, perpendicular to a fixed plane, called the scattering plane, and $F_\parallel$ is the flux density parallel to the plane. In solar system studies, polarisation is usually referred to the plane defined by the Earth, the Sun, and the object. According to (7.44), P can be positive or negative; thus the terms "positive" and "negative" polarisation are used.

The degree of polarisation as a function of the phase angle depends on the surface structure and the atmosphere. The degree of polarisation of the light reflected by the surface of an atmosphereless body is positive when the phase angle is greater than about 20°. Closer to opposition, polarisation is negative. A dependence between the polarisation and geometric albedo has been observed. This gives an independent method for determining the albedo and the size.

When light is reflected from an atmosphere, the degree of polarisation as a function of the phase angle is more complicated. For some phase angles P can be highly negative. Using the theory of radiative transfer, one can compute how the atmosphere affects light and its polarisation. Comparing these results with observations one can obtain information about the contents of the atmosphere. For example, the composition of Venus' atmosphere could be studied by polarisation studies before any probes were sent to the planet.

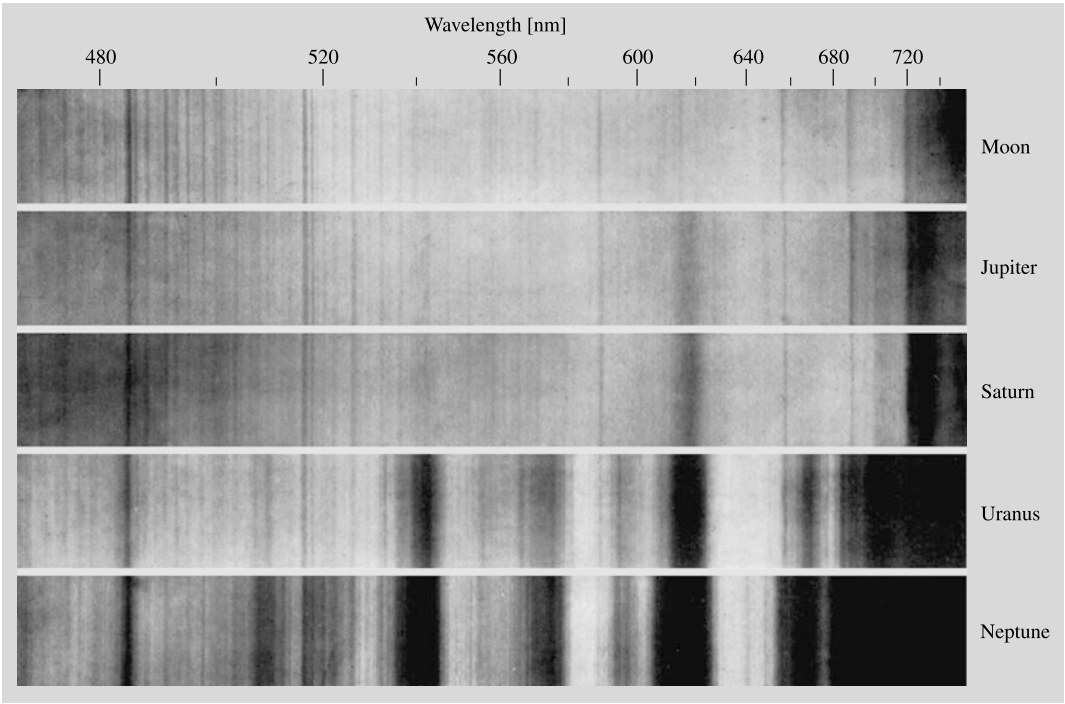


Fig. 7.23 Spectra of the Moon and the giant planets. Strong absorption bands can be seen in the spectra of Uranus and Neptune. (Lowell Observatory Bulletin **42** (1909))

Planetary Spectroscopy The photometric and polarimetric observations discussed above were monochromatic. However, the studies of the atmosphere of Venus also used spectral information. Broadband UBV photometry or polarimetry is the simplest example of spectrophotometry (spectropolarimetry). The term spectrophotometry usually means observations made with several narrowband filters. Naturally, solar system objects are also observed by means of "classical" spectroscopy.

Spectrophotometry and polarimetry give information at discrete wavelengths only. In practise, the number of points of the spectrum (or the number of filters available) is often limited to 20–30. This means that no details can be seen in the spectra. On the other hand, in ordinary spectroscopy, the limiting magnitude is smaller, although the situation is rapidly improving with the new generation detectors, such as the CCD camera.

The spectrum observed is the spectrum of the Sun. Generally, the planetary contribution is relatively small, and these differences can be seen when the solar spectrum is subtracted. The Uranian spectrum is a typical example (Fig. 7.23). There are strong absorption bands in the near-infrared. Laboratory measurements have shown that these are due to methane. A portion of the red light is also absorbed, causing the greenish colour of the planet. The general techniques of spectral observations are discussed in the context of stellar spectroscopy in Chap. 9.

7.10 Thermal Radiation of the Planets

Thermal radiation of the solar system bodies depends on the albedo and the distance from the Sun, i.e. on the amount of absorbed radiation. Internal heat is important in Jupiter and Saturn, but we may neglect it at this point.

By using the Stefan-Boltzmann law, the flux on the surface of the Sun can be expressed as

$$L = 4\pi R_{\odot}^2 \sigma T_{\odot}^4.$$

If the Bond albedo of the body is A, the fraction of the radiation absorbed by the planet is $(1 - A)$. This is later emitted as heat. If the body is at a distance r from the Sun, the absorbed flux is

$$L_{\text{abs}} = \frac{R_\odot^2 \sigma T_\odot^4 \pi R^2}{r^2}(1 - A). \tag{7.45}$$

There are good reasons to assume that the body is in *thermal equilibrium*, i.e. the emitted and the absorbed fluxes are equal. If not, the body will warm up or cool down until equilibrium is reached.

Let us first assume that the body is rotating slowly. The dark side has had time to cool down, and the thermal radiation is emitted mainly from one hemisphere. The flux emitted is

$$L_{\text{em}} = 2\pi R^2 \sigma T^4, \tag{7.46}$$

where T is the temperature of the body and $2\pi R^2$ is the area of one hemisphere. In thermal equilibrium, (7.48) and (7.49) are equal:

$$\frac{R_\odot^2 T_\odot^4}{r^2}(1 - A) = 2T^4,$$

whence

$$T = T_\odot \left(\frac{1 - A}{2}\right)^{1/4} \left(\frac{R_\odot}{r}\right)^{1/2}. \tag{7.47}$$

A body rotating quickly emits an approximately equal flux from all parts of its surface. The emitted flux is then

$$L_{\text{em}} = 4\pi R^2 \sigma T^4$$

and the temperature

$$T = T_\odot \left(\frac{1 - A}{4}\right)^{1/4} \left(\frac{R_\odot}{r}\right)^{1/2}. \tag{7.48}$$

The theoretical temperatures obtained above are not valid for most of the major planets. The main "culprits" responsible here are the atmosphere and the internal heat. Measured and theoretical temperatures of some major planets are compared in Table 7.3. Venus is an extreme example of the disagreement between theoretical and actual figures. The reason is the *greenhouse effect*: radiation is allowed to enter, but not to exit. The same effect is at work in the Earth's atmosphere. Without the greenhouse effect, the mean temperature could be well below the freezing point and the whole Earth would be ice-covered. Particularly strong the effect is on Venus, where the surface temperature is hundreds of degrees higher than the theoretical value.

Table 7.3 Theoretical and observed temperatures of some planets

	Albedo	Distance from the Sun [AU]	Theoretical temperature [K]		Observed maximum temperature [K]
			(7.50)	(7.51)	
Mercury	0.06	0.39	525	440	700
Venus	0.76	0.72	270	230	750
Earth	0.36	1.00	290	250	310
Mars	0.16	1.52	260	215	290
Jupiter	0.73	5.20	110	90	130

According to the Wien displacement law (5.22) $\lambda_{\max} = b/T$ the radiation maximum of a body at 200 K is at $\lambda = 14$ μm, deep in infrared. When the thermal radiation in the infrared or radio range is measured the temperature can be found, and further the Bond albedo can be calculated from (7.47) or (7.48). If also the phase function is known the geometric albedo and hence the diameter can be evaluated.

7.11 Origin of the Solar System

Cosmogony is a branch of astronomy which studies the origin of the solar system. The first steps of the planetary formation processes are closely connected to star formation.

Although the properties and details of the bodies of our solar system (see next chapter) may look wildly different there are some distinct features which have to be explained by any serious cosmogonical theory. These include:

- planetary orbits are almost coplanar and also parallel to the solar equator;
- orbits are almost circular;
- planets orbit the Sun counterclockwise, which is also the direction of solar rotation;

Table 7.4 True distances of the planets from the Sun and distances according to the Titius–Bode law (7.49)

Planet	n	Calculated distance [AU]	True distance [AU]
Mercury	$-\infty$	0.4	0.4
Venus	0	0.7	0.7
Earth	1	1.0	1.0
Mars	2	1.6	1.5
Ceres	3	2.8	2.8
Jupiter	4	5.2	5.2
Saturn	5	10.0	9.2
Uranus	6	19.6	19.2
Neptune	7	38.8	30.1
Pluto	8	77.2	39.5

- planets also rotate around their axes counterclockwise (excluding Venus and Uranus);
- planets have 99 % of the angular momentum of the solar system but only 0.15 % of the total mass;
- terrestrial and giant planets exhibit physical and chemical differences;
- relative abundances of ices and rocks as a function of the distance from the Sun.

Sometimes also the empirical Titius-Bode law is included (Table 7.4). It states that

$$a = 0.4 + 0.3 \times 2^n,$$
$$n = -\infty, 0, 1, 2, \ldots \qquad (7.49)$$

where the semimajor axis a is expressed in au.

It is sometimes mentioned that the first scientific theory was the vortex theory by the French philosopher *René Descartes* in 1644; however it was concerned about the motion of the solar system bodies and not its origin.

The first modern cosmogonical theories were introduced in the 18th century. One of the first cosmogonists was *Immanuel Kant*, who in 1755 presented his *nebular hypothesis*. According to this theory, the solar system condensed from a large rotating nebula. Kant's nebular hypothesis is surprisingly close to the basic ideas of modern cosmogonical models. In a similar vein, *Pierre Simon de Laplace* suggested in 1796 that the planets have formed from gas rings ejected from the equator of the collapsing Sun.

The main difficulty of the nebular hypothesis was its inability to explain the distribution of angular momentum in the solar system. Although the planets represent less than 1 % of the total mass, they possess 98 % of the angular momentum. There appeared to be no way of achieving such an unequal distribution. A second objection to the nebular hypothesis was that it provided no mechanism to form planets from the postulated gas rings.

Already in 1745, *Georges Louis Leclerc de Buffon* had proposed that the planets were formed from a vast outflow of solar material, ejected upon the impact of a large comet. Various *catastrophe theories* were popular in the 19th century and in the first decades of the 20th century when the cometary impact was replaced by a close encounter with another star. The theory was developed, e.g. by *Forest R. Moulton* (1905) and *James Jeans* (1917).

Strong tidal forces during the closest approach would tear some gas out of the Sun; this material would later accrete into planets. Such a close encounter would be an extremely rare event. Assuming a typical star density of 0.15 stars per cubic parsec and an average relative velocity of 20 km/s, only a few encounters would have taken place in the whole Galaxy during the last 510^9 years. The solar system could be a unique specimen. This is clearly against modern observations (Chap. 22).

The main objection to the collision theory is that most of the hot material torn off the Sun would be thrown out to space, rather than remaining in orbit around the Sun. There also was no obvious way how the material could form a planetary system.

In the face of the dynamical and statistical difficulties of the collision theory, the nebular hypothesis was revised and modified in the 1940's. In particular, it became clear that magnetic forces and gas outflow could efficiently transfer angular momentum from the Sun to the planetary nebula. The main principles of planetary formation are now thought to be reasonably well understood.

The oldest rocks found on the Earth are about 3.7×10^9 years old; some lunar and meteorite

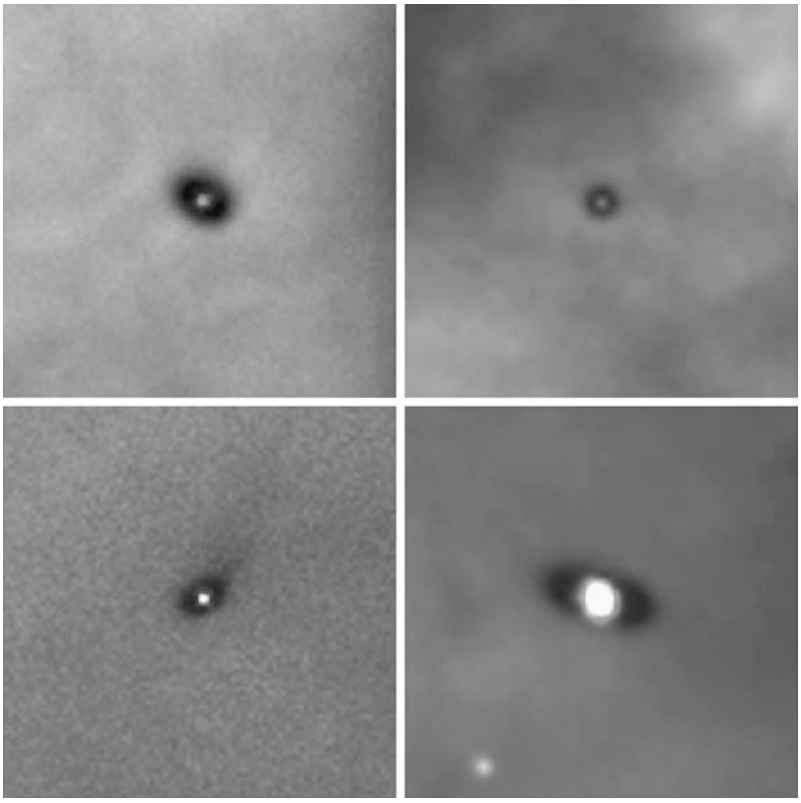


Fig. 7.24 Hubble Space Telescope images of four protoplanetary disks, "*proplyds*", around young stars in the Orion nebula. The disk diameters are two to eight times the diameter of our solar system. There is a T Tauri star in the centre of each disk. (Mark McCaughrean/Max-Planck-Institute for Astronomy, C. Robert O'Dell/Rice University, and NASA)

samples are somewhat older. When all the facts are put together, it can be estimated that the Earth and other planets were formed about 4.56×10^9 years ago. On the other hand, the age of the Galaxy is at least twice as high, so the overall conditions have not changed significantly during the lifetime of the solar system. Moreover, there is even direct evidence nowadays, such as other planetary systems and protoplanetary disks, *proplyds* (Fig. 7.24).

The Sun and practically the whole solar system simultaneously condensed from a rotating collapsing cloud of dust and gas, the density of which was some 10,000 atoms or molecules per cm^3 and the temperature 10–50 K. The elements heavier than helium were formed in the interiors of stars of preceding generations, as will be explained in Sect. 12.8. The collapse of the cloud was initiated perhaps by a shock wave emanating from a nearby supernova explosion.

The original mass of the cloud must be thousands of Solar masses to exceed the *Jeans mass*. When the cloud contracts the Jeans mass decreases. Cloud fragments and each fragment contract independently as explained in later chapters of star formation. One of the fragments became the Sun.

When the fragment continued its collapse, particles inside the cloud collided with each other. Rotation of the cloud allowed the particles to sink toward the same plane, perpendicular to the rotation axis of the cloud, but prevented them from moving toward the axis. This explains why the planetary orbits are in the same plane.

The mass of the proto-Sun was larger than the mass of the modern Sun. The flat disk in the plane of the ecliptic contained perhaps 1/10 of the total mass. Moreover, far outside, the remnants of the outer edges of the original cloud were still moving toward the centre. The Sun was losing its angular momentum to the surrounding gas by means of the magnetic field. When nuclear reactions were ignited, a strong solar wind carried away more angular momentum from the Sun.

The final result was the modern, slowly rotating Sun.

Gravitational and viscous torques transferred the angular momentum outwards. The former means a density wave caused by the instability of the disk, transferring bot mass and angular momentum outwards. Collisions between dust particles increased the velocities of outer particles and slowed down inner particles. Thus most particles moved inwards but the angular momentum outwards and the disk spread out.

Later, when nuclear reactions started, the strong solar wind transferred more angular momentum. At this *T Tauri* stage the protosun lost mass as much as $10^{-8}\ M_\odot/a$ in the form of solar wind.

Collisions of the disk particles continued. Initially individual particles stick together because of the weak intermolecular *van der Waals forces*. In less than 10,000 years the particle size increased from a few micrometres to millimetres. The growth rate was then proportional to the cross sections of the particles.

When the particles became bigger the growth rate increased considerably and became proportional to the fourth power of the particle radius. The reason for this was that the weak gravitation of bigger particles started attract gas and dust. If the mass of a particle is M and radius R and the relative velocity of a dust particle V_0 (Fig. 7.25) the effective cross section of collisions to the bigger particle is s^2:

$$s^2 = \left(R^2 + \frac{2GMR}{V_0^2}\right). \qquad (7.50)$$

Since $M \propto R^3$ we have $s^2 \propto R^4$.

The velocity of the gas was about 0.5 % smaller than the orbital velocity, and thus particles moved faster than gas and swept away the gas and dust. This resulted in rapidly growing *planetesimals*, with diameters from a few metres to kilometres.

Since big particles were moving faster than the gas they experienced a small friction slowing their velocity. The effect was strongest on metre size particles. Thus small planetesimals had to grow bigger in a few thousand years or drift down to the Sun.

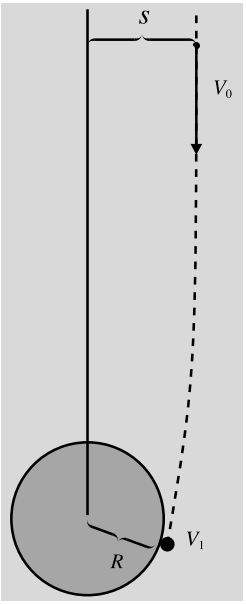


Fig. 7.25 If a particle passes a massive object at a close distance it will hit the larger body and increase its mass

When the planetesimals collided (Fig. 7.26) they grew bigger but the growth rate was no more proportional to the fourth power of the radius but slower. When the planetesimals reached the size of planets their mutual gravitation became increasingly important. Collisions of planetesimals and protoplanets shaped the solar system until it to some extent looked like the current system. The formation of the Moon, the slow retrograde of Venus and the abnormal orientation of the rotation axis of Uranus were caused by collisions of objects of the size of Mars.

The formation of Jupiter and Saturn took about 10^3–10^6 years, terrestrial planets 10^6–10^7 years, Uranus and Neptune 10^7–10^8 years. The *Nice model* (Sect. 7.12) suggests that originally Neptune was closer to the Sun than Uranus. Resonances caused Saturn, Uranus and Neptune to drift farther from the Sun, whence Neptune moved outside Uranus. Jupiter, on the other hand, moved closer to the Sun.

The strong perturbations by Jupiter prevented the formation of a large planet between Mars and Jupiter. The objects in this asteroid belt are either planetesimals or shattered protoplanets.

Depending on the volatility the matter of the solar system can be divided roughly into three categories: *Gases*, mainly hydrogen and helium, consisting of about 98.2 % of the total mass of

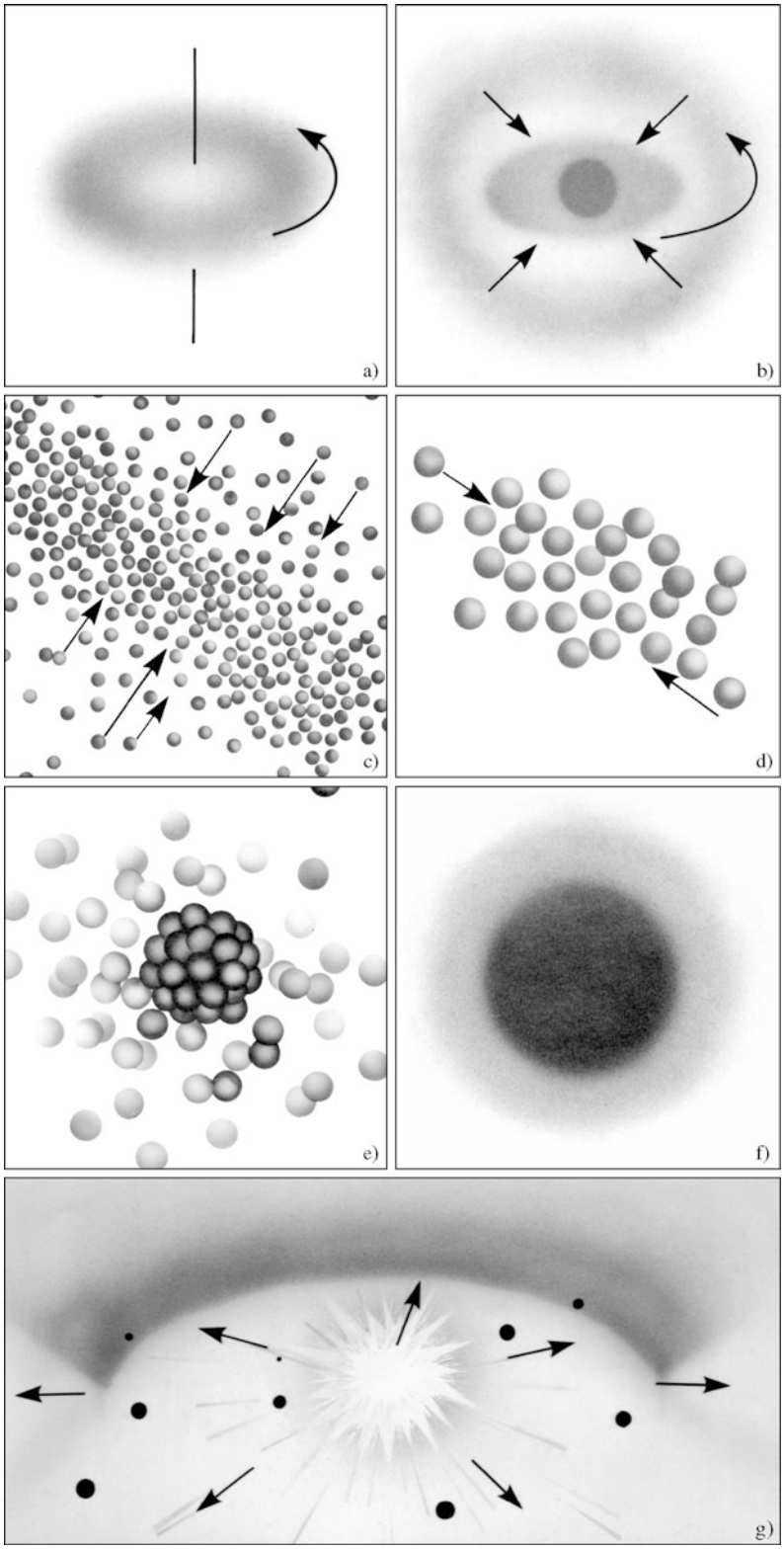


Fig. 7.26 A schematic plot on the formation of the solar system. (**a**) A large rotating cloud, the mass of which was 3–4 solar masses, began to condense. (**b**) The innermost part condensed most rapidly and a disk of gas and dust formed around the proto-sun. (**c**) Dust particles in the disk collided with each other forming larger particles and sinking rapidly to a single plane. (**d**) Particles clumped together into planetesimals which were of the size of present asteroids. (**e**) These clumps drifted together, forming planet-size bodies which began (**f**) to collect gas and dust from the surrounding cloud. (**g**) The strong solar wind "blew" away extra gas and dust; the planet formation was finished

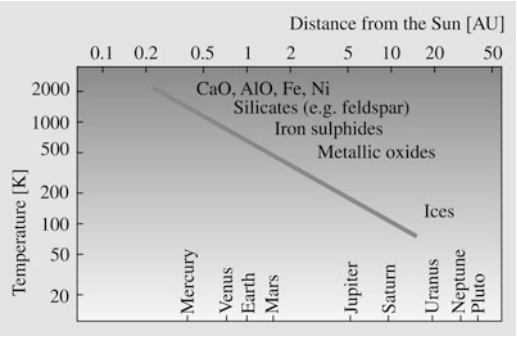


Fig. 7.27 Temperature distribution in the solar system during planet formation. The present chemical composition of the planets reflects this temperature distribution. The approximate condensing temperatures of some compounds have been indicated

Table 7.5 Mass distribution of the solar system

	Part of the (%) total mass
Sun	99.80
Jupiter	0.10
Comets	0.05
Other planets	0.04
Moons and rings	0.00005
Asteroids	0.000002
Dust	0.0000001

the solar system and remaining gaseous until very close to the absolute zero. *Ices*, about 1.4 %, melting around 160 K at the pressure of the initial nebula. *Rocks*, about 0.4 %, melting over temperatures exceeding 1000 K (Fig. 7.27).

Planets from Mercury to Mars consist mainly of rocks. When they were born the temperature in that region was too high for gases and ices to remain bound to planets. In this region over 99 % of the matter remained outside the planets. The temperature distribution is seen in the chemical contents of the planets. At the distance of Mercury the temperature had decreased below 1400 K, which meant that compounds of iron and nickel could condense from the nebula. In fact, they form about 60 % of the mass of Mercury. When we move outwards other elements become more abundant. At the distance of the Earth the temperature is about 600 K and near Mars only 450 K. The mantle of the Earth contains about 10 % of iron(II)oxide FeO. In Mars there is considerably more FeO, but in Mercury hardly anything at all.

Table 7.5 gives the mass distribution of the solar system and Table 7.6 the minimum mass needed for the existing planets. This takes into account the different composition of the planets and the Sun. In reality, the mass of the accretion disk must be much bigger, since not all of the mass did not end up in planets.

Using the minimum mass we can also calculate the required density distribution of the accretion disk. If the mass of the planet is M and it has accreted its material in the distance range ($[r_0, r_1]$) from a disk whose density is $\rho(r)$, we get

$$M = \int \rho(r)\,\mathrm{d}A = \int_0^{2\pi} \int_{r_0}^{r_1} \rho(r)\, r\,\mathrm{d}r\,\mathrm{d}\theta$$
$$= 2\pi \int_{r_0}^{r_1} \rho(r)\, r\,\mathrm{d}r. \qquad (7.51)$$

The density profile of the disk seems to obey a r^{-2} law pretty well except in the asteroid belt, where these is a clear mass defect (Fig. 7.28).

At the distance of Jupiter and Saturn the temperature was already so low that icy bodies could form. Some satellites of Saturn are examples of such bodies. From the surrounding cloud the giant planets collected gas that could stay around the planets because they were relatively far from the Sun. Jupiter and Saturn contain mostly hydrogen and helium. In Uranus and Neptune the content of these gases is smaller, possibly around twenty percent.

Continuous collisions of meteoroids, shrinking of the planets under their own gravity, and radioactive decay of relatively short lived nuclei produced a lot of heat. Heating caused partial melting of planets, leading to *differentiation*: heavier elements sank down and lighter ones rose towards the surface.

The bombardment continued for about half a billion years. Its effects are still seen on most solid bodies. For instance, the Lunar maria are remnants of that era. On the Earth the tectonic resurfacing and erosion have destroyed most meteorite craters.

Table 7.6 Minimum mass of the primordial nebula needed for the planets. The factor is a value by which the mass of the planet has to be multiplied to make the composition consistent with the Sun. The Nice model will change the values of this table and Fig. 7.28

	Distance [au]	Mass Earth = 1	Factor	Total mass	Cumulative mass
Mercury	0.4	0.055	350	19.3	19
Venus	0.7	0.815	270	220.1	239
Earth + Moon	1.0	1.012	235	237.8	477
Mars	1.5	0.107	235	25.1	502
Asteroids	2.8	0.002	200	0.4	503
Jupiter	5.2	317.89	5	1589.5	2092
Saturn	9.6	95.17	8	761.4	2853
Uranus	19.2	14.56	15	218.4	3072
Neptune	30.1	17.24	20	344.8	3417
Pluto	40	0.005	70	0.4	3417

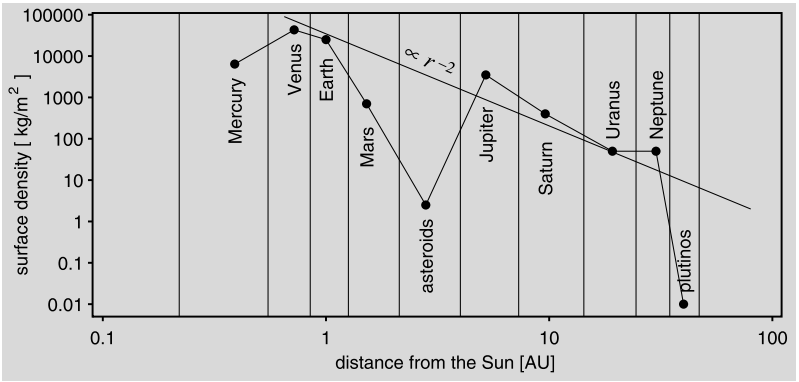


Fig. 7.28 Surface density [kg/m^2] of the accretion disc as a function of distance. The density follows approximately an r^{-2} law. Especially around the asteroid belt there seems to be a mass deficit, indicating that a considerable amount of matter has been removed elsewhere. *The vertical lines* separate regions from which each planet has accreted its material

Due to the perturbations by large planets the "leftover" planetesimals collided to planets or were thrown out to the outskirts of the solar system or even out to the interstellar space. What remained where mainly the asteroids currently on stable orbits. Lots of low-density objects, comets, were thrown to the outer regions of the solar system. These form the current *Oort cloud*. The total mass of the Oort cloud may be even 40 $M_\oplus$ and it may contain billions of comets.

Also the small bodies beyond the orbit of Neptune and the somewhat more distant *Kuiper belt* may have originated nearer to the Sun.

Planetary formation ended when the nuclear reactions of the Sun started and the Sun entered its *T Tauri stage* (Sect. 14.3). The strong solar wind caused the Sun to lose mass and angular momentum. The mass loss was about 10^{-6} $M_\odot$ a year, yet altogether maybe less than 0.1 $M_\odot$. The solar wind blew away the gas dust still in the interplanetary space, and thus the planets could not accrete any more matter.

The solar wind or radiation pressure has no effect on millimetre- and centimetre-sized particles. However, they will drift into the Sun because of the *Poynting–Robertson effect*, first introduced by John P. Poynting in 1903. Later H.P. Robertson derived the effect by using the theory of relativity. When a small body absorbs and emits radiation, it loses its orbital angular momentum and the body spirals to the Sun. At the distance of the asteroid belt, this process takes only a million years or so. Therefore the meteors wee see nowadays must be much younger than the solar system. A relatively big fraction of them is material disrupted from comets.

7.12 Nice Models

It has been assumed that the distances of the planets have not changed much since they were born. The most essential feature of the Nice model is that the giant planets were born much closer to the Sun. The model is based on a large number of computer simulations, carried out in the Côte d'Azur observatory near Nice.

The planets drifted to their current orbits due to their mutual gravity and resonances. At the same time planetary perturbations made the left-over material either to collide with planets or move outwards to the asteroid belt, Kuiper belt and Oort cloud.

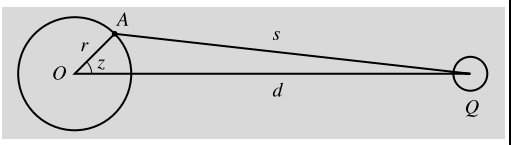
The original Nice model could not properly explain all properties of the solar system, like the structure of the Kuiper belt. These problems were fixed in the newer Nice model 2.

There are several versions of the models with varying details. They seem to approach the correct explanation, it may still be too early to discuss here their rather complicated details before they have been confirmed more convincingly.

Box 7.1 (Tides) Let the tide generating body, the mass of which is M to be at point Q at a distance d from the centre of the Earth. The potential V at the point A caused by the body Q is

$$V(A) = \frac{GM}{s}, \tag{1}$$

where s is the distance of the point A from the body Q.



Applying the cosine law in the triangle OAQ, the distance s can be expressed in terms of the other sides and the angle $z = AOQ$

$$s^2 = d^2 + r^2 - 2dr\cos z,$$

where r is the distance of the point A from the centre of the Earth. We can now rewrite (1)

$$V(A) = \frac{GM}{\sqrt{d^2 + r^2 - 2dr\cos z}}. \tag{2}$$

When the denominator is expanded into a Taylor series

$$(1+x)^{-\frac{1}{2}} \approx 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \cdots$$

where

$$x = \frac{r^2}{d^2} - 2\frac{r}{d}\cos z$$

and ignoring all terms higher than or equal to $1/d^4$ one obtains

$$V(A) = \frac{GM}{d} + \frac{GM}{d^2}r\cos z + \frac{GMr^2}{d^3}\frac{1}{2}\left(3\cos^2 z - 1\right). \tag{3}$$

The gradient of the potential $V(A)$ gives a force vector per mass unit. The first term of (3) vanishes, and the second term is a constant and independent of r. It represents the central motion. The third term of the force vector, however, depends on r. It is the main term of the tidal force. As one can see, it depends inversely on the third power of the distance d. The tidal forces are diminished very rapidly when the distance of a body increases. Therefore the tidal force caused by the Sun is less than half of that of the Moon in spite of much greater mass of the Sun.

We may rewrite the third term of (3) as

$$V_2 = 2D\left(\cos^2 z - \frac{1}{3}\right), \tag{4}$$

where

$$D = \frac{3}{4}GM\frac{r^2}{d^3}$$

is called *Doodson's tidal constant*. It's value for the Moon is 2.628 $\mathrm{m^2\,s^{-2}}$ and for the Sun 1.208 $\mathrm{m^2\,s^{-2}}$. We can approximate that z is the zenith angle of the body. The zenith angle z can be expressed in terms of the hour angle h and declination δ of the body and the latitude ϕ of the observer

$$\cos z = \cos h\cos\delta\cos\phi + \sin\delta\sin\phi.$$

Inserting this into (4) we obtain after a lengthy algebraic operation

$$V_2 = D\Big(\cos^2\phi\cos^2\delta\cos 2h + \sin 2\phi\cos 2\delta\cos h + \big(3\sin^2\phi - 1\big)\Big(\sin^2\delta - \frac{1}{3}\Big)\Big) = D(S+T+Z). \quad (5)$$

Equation (5) is the traditional basic equation of the tidal potential, the *Laplace's tidal equation*.

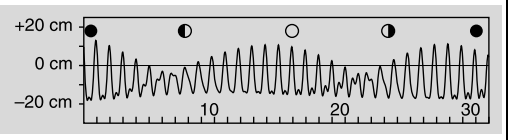
In (5) one can directly see several characteristics of tides. The term S causes the *semidiurnal tide* because it depends on $\cos 2h$. It has two daily maxima and minima, separated by 12 hours, exactly as one can obtain in following the ebb and flood. It reaches its maximum at the equator and is zero at the poles ($\cos^2\phi$).

The term T expresses the *diurnal tides* ($\cos h$). It has its maximum at the latitude $\pm 45°$ and is zero at the equator and at the poles ($\sin 2\phi$). The third term Z is independent of the rotation of the Earth. It causes the *long period tides*, the period of which is half the orbital period of the body (about 14 days in the case of the Moon and 6 months for the Sun). It is zero at the latitude $\pm 35.27°$ and has its maximum at the poles. Moreover, the time average of Z is non-zero, causing a permanent deformation of the Earth. This is called the *permanent tide*. It slightly increases the *flattening of the Earth* and it is inseparable from the flattening due to the rotation.

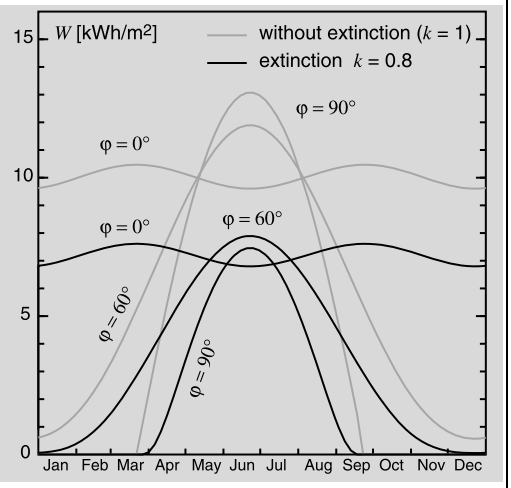
The total value of the tidal potential can be computed simply adding the potentials caused by the Moon and the Sun. Due to the tidal forces, the whole body of the Earth is deformed. The vertical motion Δr of the crust can be computed from

$$\Delta r = h\frac{V_2}{g} \approx 0.06\, V_2\ [\mathrm{m}], \quad (6)$$

where g is the mean free fall acceleration, $g \approx 9.81\ \mathrm{m\,s^{-2}}$ and h is a dimensionless number, the *Love number*, $h \approx 0.6$, which describes the elasticity of the Earth. In the picture below, one can see the vertical motion of the crust in Helsinki, Finland ($\phi = 60°, \lambda = 25°$) in January 1995. The non-zero value of the temporal mean can already be seen in this picture.



The tides have other consequences, too. Because the Earth rotates faster than the Moon orbits the Earth, the tidal bulge does not lie on the Moon–Earth line but is slightly ahead (in the direction of Earth's rotation), see below.



Due to the drag, the rotation of the Earth slows down by about 1–2 ms per century. The same reason has caused the Moon's period of rotation to slow down to its orbital period and the Moon faces the same side towards the Earth. The misaligned bulge pulls the Moon forward. The acceleration causes the increase in the semimajor axis of the Moon, about 3 cm per year.

7.13 Examples

Example 7.1 (Sidereal and Synodic Period) The time interval between two successive oppositions

of Mars is 779.9 d. Calculate the semimajor axis of Mars' orbit.

The synodic period is 779.9 d = 2.14 years. We obtain from (7.2)

$$\frac{1}{P_2} = \frac{1}{1} - \frac{1}{2.14} = 0.53 \quad \Rightarrow \quad P_2 = 1.88 \text{ a}.$$

By using Kepler's third law ($m \ll M_\odot$), the semimajor axis is found to be

$$a = P^{2/3} = 1.88^{2/3} = 1.52 \text{ au}.$$

Example 7.2 (Solar Energy Flux on the Earth) Calculate the diurnal solar energy flux per unit area at the distance of the Earth.

The solar flux density outside the Earth's atmosphere (the solar constant) is $S_0 = 1370\,\mathrm{W/m^2}$. Consider a situation at latitude ϕ, when the solar declination is δ. If the atmospheric extinction is neglected, the flux density on the surface is

$$S = S_0 \sin a,$$

where a is the elevation of the Sun. We can write $\sin a$ as a function of latitude, declination, and hour angle h:

$$\sin a = \sin\delta \sin\phi + \cos\delta \cos\phi \cos h.$$

On a cloudless day, the energy is received between sunrise and sunset. The corresponding hour angles can be obtained from the equation above, when $a = 0$:

$$\cos h_0 = -\tan\delta \tan\phi.$$

In the course of one day, the energy received on a unit area is

$$W = \int_{-h_0}^{h_0} S \,\mathrm{d}t.$$

The hour angle h is expressed in radians, so the time t is

$$t = \frac{h}{2\pi} P,$$

where $P = 1$ d $= 24$ h. The total energy is thus

$$\begin{aligned} W &= \int_{-h_0}^{h_0} S_0(\sin\delta \sin\phi + \cos\delta \cos\phi \cos h) \\ &\quad \times \frac{P}{2\pi} \,\mathrm{d}h \\ &= \frac{S_0 P}{\pi}(h_0 \sin\delta \sin\phi + \cos\delta \cos\phi \sin h_0), \end{aligned}$$

where

$$h_0 = \arccos(-\tan\delta \tan\phi).$$

For example near the equator ($\phi = 0°$) $\cos h_0 = 0$ and

$$W(\phi = 0°) = \frac{S_0 P}{\pi} \cos\delta.$$

At those latitudes where the Sun will not set, $h_0 = \pi$ and

$$W_{\mathrm{circ}} = S_0 P \sin\delta \sin\phi.$$

Near the poles, the Sun is always circumpolar when above the horizon, and so

$$W(\phi = 90°) = S_0 P \sin\delta.$$

Interestingly enough, during the summer when the declination of the Sun is large, the polar areas receive more energy than the areas close to the equator. This is true when

$$\begin{aligned} & W(\phi = 90°) > W(\phi = 0°) \\ \Leftrightarrow \quad & S_0 P \sin\delta > S_0 P \cos\delta/\pi \\ \Leftrightarrow \quad & \tan\delta > 1/\pi \\ \Leftrightarrow \quad & \delta > 17.7°. \end{aligned}$$

The declination of the Sun is greater than this about two months every summer.

However, atmospheric extinction diminishes these values, and the loss is at its greatest at the poles, where the elevation of the Sun is always relatively small. Radiation must penetrate thick layers of the atmosphere and the path length is comparable to $1/\sin a$. If it is assumed that the fraction k of the flux density reaches the surface when the Sun is at zenith, the flux density when the Sun is at the elevation a is

$$S' = S_0 \sin a \, k^{1/\sin a}.$$

The total energy received during one day is thus

$$W = \int_{-h_0}^{h_0} S' \,\mathrm{d}t = \int_{-h_0}^{h_0} S_0 \sin a \; k^{1/\sin a} \,\mathrm{d}t.$$

This cannot be solved in a closed form and numerical methods must be used.

The figure on next page shows the daily received energy W [kW h/m^2] during a year at latitudes $\phi = 0°$, $60°$, and $90°$ without extinction, and when $k = 0.8$, which is close to the real value.

Example 7.3 (Magnitude of a Planet) The apparent magnitude of Mars during the 1975 opposition was $m_1 = -1.6$ and the distance to the Sun, $r_1 = 1.55$ au. During the 1982 opposition, the distance was $r_2 = 1.64$ au. Calculate the apparent magnitude in the 1982 opposition.

At opposition, the distance of Mars from the Earth is $\Delta = r - 1$. The observed flux density depends on the distances to the Earth and the Sun,

$$F \propto \frac{1}{r^2 \Delta^2}.$$

Using the magnitude formula (4.9) we obtain

$$m_1 - m_2 = -2.5 \lg \frac{r_2^2 (r_2 - 1)^2}{r_1^2 (r_1 - 1)^2}$$

$$\Rightarrow \quad m_2 = m_1 + 5 \lg \frac{r_2 (r_2 - 1)}{r_1 (r_1 - 1)}$$

$$= -1.6 + 5 \lg \frac{1.64 \times 0.64}{1.55 \times 0.55} \approx -1.1.$$

The same result is obtained if (7.38) is separately written for both oppositions.

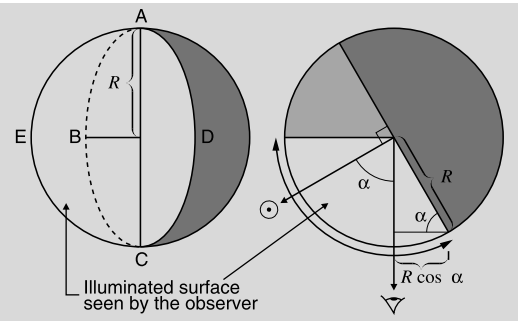
Example 7.4 (The Brightness of Venus) Find the instant when Venus is brightest if the brightness is proportional to the projected size of the illuminated surface. The orbits are assumed to be circular.

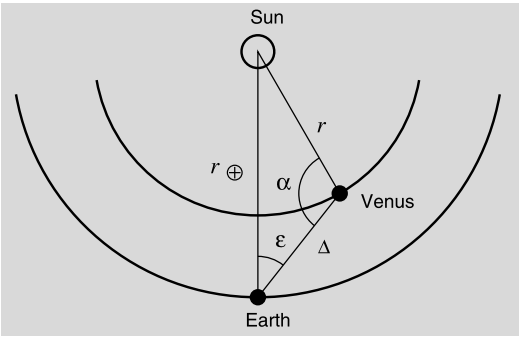
The size of the illuminated surface is the area of the semicircle $ACE\pm$ half the area of the ellipse $ABCD$. The semiaxes of the ellipse are R and $R\cos\alpha$. If the radius of the planet is R, the illuminated area is

$$\pi \frac{R^2}{2} + \frac{1}{2} \pi R \times R \cos\alpha = \frac{\pi}{2} R^2 (1 + \cos\alpha),$$

where α is the phase angle. The flux density is inversely proportional to the square of the distance Δ. Thus

$$F \propto \frac{1 + \cos\alpha}{\Delta^2}.$$





The cosine formula yields

$$M_\oplus^2 = r^2 + \Delta^2 - 2\Delta r \cos\alpha.$$

When $\cos\alpha$ is solved and inserted in the flux density we obtain

$$F \propto \frac{2\Delta r + r^2 + \Delta^2 - M_\oplus^2}{2 r \Delta^3}.$$

The minimum of the equation yields the distance where Venus is brightest:

$$\frac{\partial F}{\partial \Delta} = -\frac{4r\Delta + 3r^2 - 3M_\oplus^2 + \Delta^2}{2r\Delta^4} = 0$$

$$\Rightarrow \quad \Delta = -2r \pm \sqrt{r^2 + 3M_\oplus^2}.$$

If $r = 0.723$ au and $R_\oplus = 1$ au, the distance is $\Delta = 0.43$ au and the corresponding phase angle is $\alpha = 118°$.

Thus Venus is brightest shortly after the largest eastern elongation and before the largest western elongation. From the sine formula we obtain

$$\frac{\sin\varepsilon}{r} = \frac{\sin\alpha}{M_\oplus}.$$

The corresponding elongation is $\varepsilon = 40°$, and

$$\frac{1+\cos\alpha}{2} \times 100\,\% = 27\,\%$$

of the surface is seen lit.

Example 7.5 (Magnitude of an Asteroid) The parameters of the asteroid 44 Nysa are $H = 6.929$, $G_1 = 0.050$, $G_2 = 0.67$ and the semimajor axis of the orbit $a = 2.42$ au. Find $V(1, 1°)$. What is the apparent magnitude at phase angles 0° and 1°?

Values of the basis functions are $\Phi_1(1°) = 0.9667$, $\Phi_2(1°) = 0.9900$, $\Phi_3(1°) = 0.577$. Using these we get the absolute magnitude at the phase angle 1°:

$$\begin{aligned} V(1, 1°) &= 6.929 - 2.5\lg\big[0.050 \times 0.9667 \\ &\quad + 0.67 \times 0.990 \\ &\quad + (1 - 0.050 - 0.67) \times 0.577\big] \\ &= 7.076. \end{aligned} \tag{7.52}$$

Near the opposition we can approximate $\Delta = r - 1 = 1.42$ au. The apparent magnitude at the opposition is then

$$\begin{aligned} m &= 6.929 + 5\lg(1.42 \times 2.42) - 2.5\lg\Phi(0°) \\ &= 6.929 + 5\lg 3.36 = 9.561. \end{aligned} \tag{7.53}$$

When the phase angle is 1° the absolute magnitude is 0.147 greater than at the opposition. The apparent magnitudes differ by the same amount; hence $m(\alpha = 1°) = 9.71$.

Example 7.6 Find the distance of a comet from the Sun when its temperature reaches 0 °C and 100 °C. Assume the Bond albedo of the comet is 0.05.

Solve r from (7.47):

$$r = \left(\frac{T_\odot}{T}\right)^2 \left(\frac{1-A}{2}\right)^{1/2} R_\odot.$$

If $T = 273$ K, we the distance is $r = 1.4$ au; when $T = 373$ K, we get $r = 0.8$ au.

7.14 Exercises

Exercise 7.1 What is the greatest possible elongation of Mercury, Venus and Mars? How long before sunrise or after sunset is the planet visible? Assume that the declination of the planet and the Sun is $\delta = 0°$.

Exercise 7.2 (a) What is the greatest possible geocentric latitude of Venus, i.e. how far from the Sun can the planet be at the inferior conjunction? Assume the orbits are circular.

(b) When is the situation possible? The longitude of the ascending node of Venus is 77°.

Exercise 7.3 (a) Find the daily retrograde apparent motion of an exterior planet at its opposition. Assume that the planet and the Earth have circular orbits.

(b) Pluto was found in 1930 from two plates, exposed 6 days apart during the opposition of the planet. On those plates one degree corresponded to 3 cm. How much (in cm) had Pluto moved between the exposures? How much does a typical main belt asteroid move in the same time?

Exercise 7.4 A planet is observed at the opposition or inferior conjunction. Due to the finite speed of light the apparent direction of the planet differs from the true place. Find this difference as a function of the radius of the orbit. You can assume the orbits are circular. Which planet has the largest deviation?

Exercise 7.5 The angular diameter of the Moon is 0.5°. The full moon has an apparent magnitude of −12.5 and the Sun −26.7. Find the geometric and Bond albedos of the Moon, assuming that the reflected light is isotropic (into a solid angle 2π sterad).

Exercise 7.6 The eccentricity of the orbit of Mercury is 0.206. How much does the apparent magnitude of the Sun vary as seen from Mercury? How does the surface brightness of the Sun vary?

Exercise 7.7 An asteroid with a diameter of 100 m approaches the Earth at a velocity of 30 $\mathrm{km\,s^{-1}}$. Find the apparent magnitude of the asteroid (a) one week, (b) one day before the collision. Assume that the phase angle is $\alpha = 0°$ and the geometric albedo of the asteroid is $p = 0.1$. What do you think about the chances of finding the asteroid well in advance the crash?

Exercise 7.8 Find the centripetal acceleration at the poles and on the equator of the Earth.