Physical Processes in the Solar System

An introduction to the physics of asteroids, comets, moons and planets

First Edition

John D. Landstreet
University of Western Ontario
London, Canada

Keenan & Darlington, Publishers London, Canada

PUBLISHED BY KEENAN & DARLINGTON, PUBLISHERS,
18 Rollingwood Circle, London, Ontario, Canada N6G 1P7

KEENAN & DARLINGTON, PUBLISHERS
18 Rollingwood Circle, London, Ontario, Canada N6G 1P7
e-mail: planets@corcaroli.astro.uwo.ca
Web: http://www.astro.uwo.ca/~jlandstr/planets

First published 2003

Printed in Canada

This book was prepared with the LaTeX document preparation system.

National Library of Canada Cataloguing in Publication

Landstreet, John D., 1940-
Physical processes in the solar system: an introduction to the origin, structure and evolution of asteroids, comets, planets and moons / John D. Landstreet.

Includes bibliographical references and index.
ISBN 0–9732051–0–5

1. Solar system – – Textbooks. I. Title. QB 501.L35 2003 523.2 C2003–903439–9

ISBN 0–9732051–0–5 paperback

Chapter 11

Giant Planets and their Moons

11.1 Overview

So far we have mainly discussed the terrestrial planets, which resemble more or less closely our own Earth. Because we know quite a lot about the Earth, it has been possible to use it as a prototype in our efforts to understand the other rocky planets of the inner solar system. Now we turn our attention to the giant planets, a group of bodies which are quite different structurally from the Earth, and to their moons, bodies not as unlike the Earth as the giant planets themselves are, but having nevertheless some quite distinctive features not found in the planets and moons of the inner solar system.

The most basic data concerning the four large planets Jupiter, Saturn, Uranus, and Neptune immediately reveals that these are bodies which are quite different from the four terrestrial planets, and justifies grouping them together as the giant planets. (Pluto, the ninth planet, is a body more like the moons of the giants than like any of the other outer planets, and it may be related to the Kuiper belt of comet nuclei. It will be discussed at the end of the chapter.) In contrast to the inner planets, which range in mass from the Earth's mass on down to only about 5.5% of that mass in Mercury, the giant planets are all substantially more massive than the Earth (see Table A.2), ranging from about 15 Earth masses for Uranus and Neptune up to nearly 100 Earth masses for Saturn and over 300 for Jupiter. More than 99.5% of the total planetary mass is found in the four giant planets, more than 70% of it in Jupiter alone. Furthermore, all four giant planets are substantially larger in size than any of the terrestrial planets; the giants range from about 4 Earth radii (Neptune and Uranus) up to more than 11 (Jupiter), while the four terrestrial planets range between about 0.4 and 1.0 Earth radii in size.

Remarkably, in spite of their large sizes, the giant planets are all much less dense than the terrestrial ones. All have densities not very different from that of water (their densities range from 690 to 1640 $\mathrm{kg\,m^{-3}}$). These densities are several times smaller than those of any of the inner planets (where densities range from 3930 to 5520 $\mathrm{kg\,m^{-3}}$). It is this fact, taken together with their large sizes, that clearly signals the major difference in chemistry or structure of the giants as compared to the inner planets. Considering that the inner planets are already massive enough that their densities are increased (over the densities they would have if they were, say, only as massive as Ceres, or the Moon) by the effects of gravity in compressing their matter, it is clear that the material making up the giants cannot be the same as that of the terrestrial planets, even approximately, or densities of well over 6000 $\mathrm{kg\,m^{-3}}$ would be found. We shall see later that although all the giant planets have rock (and ice?) cores with masses of the order of 10 – 15 Earth masses, all have massive and deep atmospheres composed mainly of H and He. These atmospheres constitute an important fraction of the total mass of each giant planet, ranging from 10 or 20% of the total for Uranus and Neptune, up to 80% of Saturn and 95% of Jupiter. Such massive atmospheres are completely different from the relatively insignificant atmospheres of the terrestrial planets, which (including all the volatiles, such as water, CO_2, and N_2) constitute no more than 0.01% of total planetary mass.

The four giant planets differ in several other fundamental ways from the terrestrial planets. All but Uranus transport enough internal heat to the surface that they radiate into space roughly twice as much heat as they absorb from incident sunlight. This heat loss from the interior seems to be either radiation of heat released in the interior of the planet from gravitational contraction as it formed, or heat released by gravitational separation of He from H in the planetary interior. This is quite different from the behaviour of the rocky planets, whose internal heat loss is several orders of magnitude smaller than the absorbed and reradiated solar energy, and is due largely to the loss of heat released by radioactive decay of U, Th, and especially ^{40}K. In contrast to the scarcity of moons in the inner solar system, each giant planet has at least one large moon; in fact, each has a system of numerous moons ranging in size from the size of the Earth's Moon on down to a few km in radius. A couple of the moons of Jupiter have densities comparable to that of the Earth's Moon, but most have sufficiently low density that it is clear that they must contain a considerable amount of ice in their bulk composition. All four giants also have ring systems, although only that of Saturn is easily visible from Earth. Thus we are clearly justified in treating both the giant planets and their moons as objects which are really different from the largest bodies in the inner solar system.

We first consider the information available about these remarkable objects from observations, and then look at what is understood about their internal structure, and how they may have formed and developed during the history of the solar system.

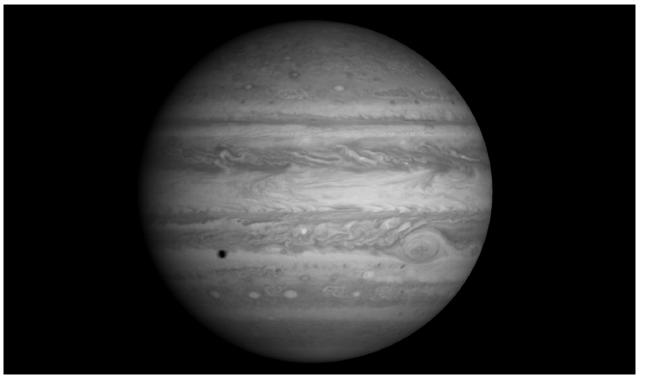

Figure 11.1: This image of Jupiter was taken by the Hubble Space Telescope. It clearly shows the complex cloud bands, alternating between light-coloured zones and darker belts. The wave-like and oval shapes in the belts are cloud patterns that reflect complex wind systems in the cloudy atmosphere of the planet. The large oval just below the equator, near the right limb, is the Great Red Spot. The small black dot at the same latitude near the left limb is the shadow of one of the moons of Jupiter.) (Courtesy of NASA.)

11.2 The giant planets

Observational data on the giant planets

What we see when we look at one of the giant planets is very different from the appearance of the terrestrial planets (except for cloud-covered Venus). The visible "surfaces" of the giants are in fact simply the uppermost of several decks of clouds, seen more or less clearly through a light haze high in the atmosphere. In Jupiter and Saturn, the upper cloud layer is composed primarily of ammonia (NH_3) ice crystals. In places it is possible to see down to the next cloud deck of ammonium hydrosulfide (NH_4SH) ices. The atmospheres of Uranus and Neptune are colder than those of Jupiter and Saturn, and so the highest cloud decks, which form the base of the visible atmospheres, are composed of ice crystals of methane (CH_4), and below that, of hydrogen sulfide (H_2S). (Nowhere in the atmospheres of Jupiter and Saturn does the temperature drop to a low enough value for methane to freeze out, hence there is no methane cloud deck on either planet.)

The rather changeable features seen on the "surfaces" of the giant planets are thus mostly due to forms and colorations of the highest cloud layers. Jupiter (see Figure 11.1) displays a disk with alternating white and coloured stripes (called respectively **zones** and **belts**) across its disk parallel to the planet's equator; in fact, these are wide bands of clouds, variously white, tan, light yellow, brown, and even red, that circle the planet at different latitudes. These cloud bands correspond to wind streams, like the jet streams of the Earth's atmosphere, in the planet's upper atmosphere. Near the equator, these winds circle the planet in the direction of its rotation (from west to east) at a speed of about 300 $\mathrm{km\,s^{-1}}$ (relative to the rotation of the planet's deep interior). There are two other regions of high speed winds near $+25°$ and $-25°$ latitude. At other latitudes the winds blow less quickly. These various wind streams seem to be alternately regions of rising and of sinking gases. The cloud tops in the rising gases occur at higher altitude, where the clouds are white ammonia crystals; the sinking gas forms clouds a little lower, where the higher temperature leads to coloured ammonium hydrosulphide ice, thus producing the banded appearance of the planet.

The other prominent feature seen on Jupiter is the **Great Red Spot**, a huge circular storm rather like a terrestrial hurricane. This storm has varied in size (it is presently about twice the size of the Earth) and in the intensity of its red colour through the years, but has been continuously present since it was discovered by the British scientist Robert Hooke (1635 – 1703), a contemporary and competitor of Issac Newton, in the mid-seventeenth century, more than 300 years ago. A number of smaller white circular storms, considerably smaller than the Red Spot, are also present on the planet at any one time, but these are not as long-lived as the Red Spot.

Saturn also presents a banded appearance like that of Jupiter, but much less vivid (Figure 11.2). The cloud colour on the second largest planet is generally yellow, but varies from nearly white to light brown. The pattern of bands is similar to that on Jupiter. There is an equatorial jet stream that flows from west to east around the planet, but the speed of this flow, about 1500 $\mathrm{km\,s^{-1}}$, is much higher than is found on Jupiter.

Figure 11.2: This image of Saturn, taken by one of the Voyager spacecraft as it approached the giant planet, shows the spectacular ring system as well as faint bands encircling the planet, like the belts and zones of Jupiter. These faint markings reveal the wind systems in the planet's deep atmosphere. (Courtesy of NASA.)

Small circular storms, like the smaller white ovals on Jupiter, are sometimes seen on Saturn, but they are not nearly as striking as the Great Red Spot, nor as long-lived.

Uranus has a generally blue-green colour. This is caused by methane in the upper atmosphere of the planet: as sunlight flows into the atmosphere where it is reflected back out by haze and the highest cloud layer, the red light is absorbed by the methane, leaving the reflected light rich in blue and green light, and deficient in red. Clouds form only at rather great depths in this cold atmosphere, and are generally obscured by the upper atmosphere haze or smog. Hence almost no cloud features or structure are seen (Figure 11.3. A very striking peculiarity of Uranus compared to all the other planets is that its axis of rotation lies almost in the plane of the ecliptic, so that as the planet circles the Sun, first one hemisphere and then the other passes through a period of complete darkness when the opposite pole is pointing towards the Sun.

Neptune has an even stronger blue colour than Uranus, also due to absorption of red light by methane, but its clouds lie at a somewhat higher altitude and are more easily seen. Like Jupiter, Neptune exhibits giant cyclonic storms. A huge storm, about as large as the Earth, was observed on the planet as the Voyager 2 spacecraft passed through the system. This storm was a darker colour than the planetary disk generally, and is known as the **Great Dark Spot**. White clouds are also seen at a few locations in the atmosphere. Both features are visible in Figure 11.4. The high-altitude winds on Neptune generally circle the planet more slowly than it rotates (that is, they blow from east to west).

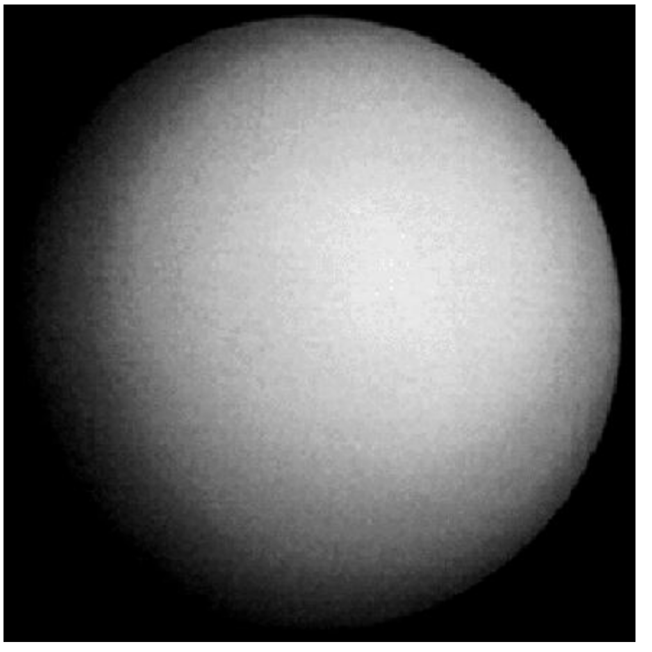

Figure 11.3: The deep atmosphere of Uranus shows no visible features at all except for a slight variation in brightness over the surface. Absorption of red light by methane in the atmosphere gives the planet a deep blue tint. (Courtesy of NASA.)

As we will discuss later, it is clear that the interiors of Jupiter, Saturn, and Neptune are convecting (this also is assumed to be true of Uranus although we have no strong direct evidence one way or the other). This convection appears to extend up to the visible layers of the atmospheres of the giants. Since convection involves vigorous mixing motions, we could wonder if this means that the atmospheres of the giants have chemical compositions that directly reflect the interior (bulk) compositions of the planets. The answer is: *no*.

The effect that drastically limits the chemical species present in the atmosphere is condensation. In the convecting interior of a giant planet, blobs and streams of gas rise while others descend. In a rising blob of gas, the temperature of the blob will decrease steadily as it rises. This temperature drop is caused by the expansion of the blob, which expends internal energy by pushing its surroundings away. As the temperature decreases in the rising blob, it will reach the temperature at which specific atoms or molecules freeze out, as we have discussed earlier in connection with the cooling solar nebula in which the planets formed. Thus refractory materials such as iron and silicate rock will

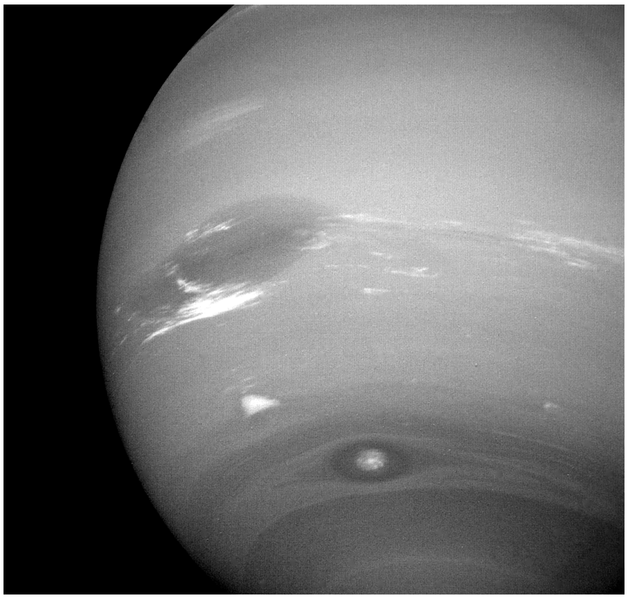

Figure 11.4: This close-up view of the disk of the planet Neptune was obtained by the Voyager 2 spacecraft as it passed Neptune in 1989. Near the middle of the image is a huge cyclonic storm known as the Great Dark Spot. It is roughly a large as Jupiter's Great Red Spot. Below the Great Dark Spot is a smaller storm. Several patches of white cloud are visible. (Courtesy of NASA.)

condense where the temperature is above 2000 K, 1000 km or more below the surface.

The ices will form liquid drops or solid flakes much nearer the surface. In a rising stream of gas in Jupiter's interior, for example, we expect NH_4Cl to condense at about 100 km below the visible surface, water (perhaps mixed with ammonia) will condense at about 50 km below the surface, and NH_4SH at about 20 km down. NH_3 will condense in the visible layers. At the level at which an abundant substance condenses, the droplets or solid particles formed will produce clouds. Within these clouds, the updraft will continue to lift the drops or particles as they grow by further condensation, but when the cloud particles become large enough to not be supported by the rising gases, they will fall back through the cloud layer into the warmer gas below, where they will evaporate again.

We therefore expect that there should be a series of major cloud decks below the visible atmosphere of each of the giant planets. In Jupiter and Saturn, it appears that the highest deck that is really opaque is that of NH_4SH. When we study the chemistry of the atmosphere, we are mainly limited to the layers above this level. Because of condensation of most elements below this cloud deck, the only elements that are expected to have the same relative abundances in the atmosphere as they do deep in the planet are H, C (since the predominant compound of C, CH_4, condenses only at a temperature lower than that reached anywhere in the atmospheres of Jupiter or Saturn), and the noble gases He, Ne, Ar, Kr and Xe (all of which condense only at very low temperatures). Essentially all the other elements, including the abundant and important substances N (in NH_3) and O (in H_2O), will be greatly depleted in the observable layers relative to their contribution to the bulk chemical composition.

We have obtained information about the elements that are present in the atmospheres of the giant planets by a number of different methods. The earliest technique used was to look at the spectrum of reflected sunlight at visible or near infrared wavelengths, and see what additional colours are weakened compared to the directly observed solar spectrum. Such observations may be done from the ground, and reveal the presence of CH_4, some NH_3, and H_2. The possibilities of remote sensing (measurements done without benefit of expensive trips to the outer planets) have been greatly expanded in recent years by the availability of spectrographs on satellites which were used to study reflected ultraviolet sunlight, and the thermal radiation emitted directly by the warm atmospheres of the giants beyond 5 μm in wavelength. Such observations add water, phosphine (PH_3) and several trace hydrocarbons such as C_2H_6 to the observed substances.

The study of the giant planets (and their moons) was completely revolutionized by the data sent back by the two unmanned Voyager spacecraft, and the more recent Galileo probe. Both Voyager spacecraft passed through the Jupiter system in 1979, and visited Saturn in 1980 (Voyager 1) and 1981 (Voyager 2). Voyager 1 passed close to Saturn's largest moon, Titan, and subsequently continued on a trajectory out of the plane of the ecliptic. Voyager 2 continued on in the ecliptic and passed close to Uranus in 1986 and Neptune in 1989. In 1996 the Galileo mission placed a large satellite into orbit around Jupiter, and dropped a probe into the atmosphere of the giant planet. And in 2004 the Cassini space probe is scheduled to go into orbit around Saturn. These missions have made it possible to obtain spectacular close-up views of all the outer planets except Pluto. They have provided us with fascinating images of surfaces and atmospheres as well as with a variety of other measurements such as spectroscopic observations of small surface regions.

The Voyagers and Galileo all carry spectrographs of various kinds that have supplemented in extremely valuable ways the spectra that can be obtained from Earth. But perhaps the most interesting result about chemistry from these missions came from the Galileo project, which included a small probe that actually entered the atmosphere of Jupiter and directly sampled the composition of the gases there, down to layers well below those directly accessible from the outside. The probe was also instrumented to report when it passed though cloud layers. Unfortunately, the probe apparently entered Jupiter in a small region of relatively clear gas (perhaps a column of downflowing gas that had been cleared of most condensible substances), so that the thin cloud layers that it passed through were hardly representative of the rest of the atmosphere.

The results of the observations, both remote and from space probes, are as follows. The main gas present in the atmospheres of all four giant planets is hydrogen, in the form of H_2 molecules, which are detected directly by their absorption lines in the near infrared portion of the spectrum, around 8000 Å. The next most important atomic species in the atmospheres is He, which (because it simply does not form molecules at all) is in atomic form. These two elements account for almost all the atoms in the atmospheres of all the giant planets; all other species (C, N, O, etc) are simply trace elements, present at the level of one or two atoms in a hundred, or less. Cold He cannot be detected in any planetary spectrum, because it has no spectral lines anywhere in the visible or infrared. Its presence and abundance are deduced from the radio occultation experiments carried out by the Voyager and Galileo spacecraft. In these experiments, a pure radio tone (frequency) broadcast by the spacecraft as it passed behind the planet, and again as it emerged on the other limb, was observed from Earth. The refraction (bending) of the radio beam by the planetary atmosphere (which made the spacecraft seem to move at a different speed along its path in space than it was actually moving) was measured. The information about refraction was used to determine the average molecular weight of the atmospheric gas. Since only H and He contribute appreciably to this average molecular weight, the ratio of numbers of He to H atoms may be determined. It is found that for Jupiter, Uranus, and Neptune, He makes up about the same fraction of the total atmospheric gas that it does in the Sun, where there is about one He atom for every ten H atoms (see Table 2.2), so that He makes up about 30% of the total atmospheric mass. However, for Saturn, it appears that He is depleted relative to this value by about a factor of two; only 6% of the gas atoms in the atmosphere are He. The explanation for this interesting fact seems to be the partial separation of H and He in the deep interior of Saturn, as will be discussed in a later section.

The other elements for which abundances might be determined from the spectrum of reflected sunlight are the abundant light elements C, N, and O. (Neon, about as abundant as nitrogen in the sun, has no spectral lines in the visible or infrared spectrum.) The three detectable light elements are almost entirely bound up in the molecules H_2O, CH_4, and NH_3, so one needs to study the spectra of these molecules to derive the abundances of C, N, and O. Determination of the O abundance is quite difficult for all the giant planets because H_2O freezes out as clouds deep in the atmosphere, so that the number of water molecules in the upper atmosphere is greatly depleted. The formation of clouds of ammonia or of ammonium hydrosulphide similarly limits the number of NH_3 molecules in the upper atmosphere, but the abundance of N can be estimated in the atmospheres of Jupiter and Saturn, where it appears to be similar to, or slightly larger than, the abundance found in the Sun. The large abundance of N is confirmed for Jupiter by the Galileo probe, which found that N settled to about three times its solar abundance deep in the atmosphere. The results for Uranus are much less certain, and significant variations in abundance from one latitude to another seem to be found, but it appears that N is somewhat underabundant compared to the Sun. For Neptune, no clear result is available yet.

The best determined of the abundant light elements in all four giant planet atmospheres is C, because methane is sufficiently volatile that it condenses (if at all) only in a high, easily observed region of the atmosphere. C appears to be mildly more abundant in Jupiter and Saturn than in the Sun, by about a factor of three in Jupiter, and six in Saturn. The results for Uranus and Neptune are less consistent, but suggest an overabundance of C, relative to the solar C fraction, of perhaps as much as 20. This may well reflect some mixing up into the atmosphere of the ices that make up a large fraction of the planetary mass of these two bodies.

The heavier noble gases Ar, Kr and Xe appear to be present in Jupiter's atmosphere at about 2 − 3 times their abundances in the sun.

Several characteristics of the giant planets, in addition to those that may be deduced from their atmospheric appearances or spectra, can help us to understand their origins and evolution. One particularly interesting feature is the remarkable difference between the complex moon systems found in the outer solar system, and the comparative rarity of moons around the terrestrial planets. In the inner solar system, only the Earth has a significant (but rather anomalous) moon; Mars has two tiny satellites which probably are left from the period when the planet accreted. The giants, on the other hand, have more than 50 moons among them. These moons form highly ordered systems around Jupiter and Saturn.

Jupiter has four moons comparable in size to the Earth's Moon, which orbit the planet in nearly circular, prograde (i.e. in the direction of the planet's rotation) orbits in the planet's equatorial plane. At least four other small moons also orbit Jupiter in similar orbits. However, the planet also has at least eight moons that orbit it in much larger, inclined and eccentric orbits, some of which are retrograde. These two sets of moons, which we call **regular** and **irregular**, have such different orbital characteristics that it is usually assumed that they must have quite different origins. Saturn also has a large system of regular satellites, at least 18 of them, including one (Titan) which is considerably larger than the Earth's Moon. Saturn has two known irregular moons. Uranus has at least 15 regular moons and two irregular ones, although none are more than about 800 km in radius. Neptune has one large moon (in a retrograde orbit, however!) as well as at least 6 smaller regular moons in small, circular orbits, and one other irregular moon in a very much larger orbit. All four giant planets thus have very similar moon systems, including one or more quite large satellites, a number of moons in regular orbits, and at least one in an irregular orbit. The fact that a large number of moons is found around each of the giant planets strongly suggests that these satellite systems formed as a normal part of the process of planet building, and we shall see that the moon systems contain important clues about giant planet formation.

The giant planets are also all now known to have ring systems. The spectacular rings of Saturn, easily visible in a small telescope (Figure 11.2), have been known since the seventeenth century. They were first seen (quite imperfectly) in 1610 by Galilei, who thought that perhaps Saturn was very non-spherical, or possibly even made up of several bodies close to one another. The nature of the Saturnian disk was first understood in 1659 by the Dutch astronomer Christiaan Huygens. In 1675 the French-Italian astronomer Giovanni Domenico Cassini (1625 – 1712) observed that the ring, which to the casual observer seems to be a continuous, almost solid band (like the brim of a hat, although not attached to the planet), actually has a dark gap about two-thirds of the way out, which divides the ring into a broad inner section and a narrower outer section. This division is clearly visible in Figure 11.2. More recently, other astronomers have observed from Earth that the ring actually shows several dark dividing lanes. Close-up pictures from the two Voyager spacecraft, taken as they passed through the Saturn system in 1980 and 1981, reveal that the rings are actually made up of literally thousands of narrow ringlets, each defining a narrow band around Saturn's equator which is relatively full of the ring material, and separated from neighboring ringlets by equally narrow regions with less material. The distribution of material around each of the ringlets varies in time, often in synchronism with other ringlets, giving rise to patterns like spokes in a wheel.

Rings about the other three giants were only discovered recently. Those of Uranus were observed by chance from Earth in 1977, when astronomers observing the occultation (eclipse) of a distant star by Uranus itself saw the starlight briefly dimmed several times both before and after the planetary occultation, as nine thin rings passed in front of the star. Two more rings were discovered by Voyager 2. Unlike the rings of Saturn, those of Uranus are narrow, dark, and widely separated. One faint, broad, and diffuse ring of dark material was discovered around Jupiter in 1979 by the Voyagers. Earth-based occultation measurements and observations from Voyager 2 show that Neptune has four dark rings, one broad and diffuse like that of Jupiter, and three narrow rings similar to those of Uranus. The rings of Jupiter, Uranus, and Neptune differ from those of Saturn in containing far less material, and in being made of dark material (like rock) rather than bright (like ice).

Composition of the giant planets

We next try to understand the general structure of the giant planets. We shall want especially to determine, as far as we can, what their bulk chemical compositions are, whether they are layered in some way, as the Earth is, and how hot they are inside. The giants are very different in size, in average density, and in general chemical composition from the Earth, and we cannot use the Earth as an approximate starting point as we did for the terrestrial planets. Instead, we must consider a larger range of possibilities.

First we need to see what kinds of material might have gone into the construction of the giant planets. We could imagine trying to somehow determine the chemical makeup element by element, but it is not hard to see that this will give us a problem with so many unknown quantities that we have no hope of finding a unique solution. Instead, we need to pick a small number of plausible chemical substances or mixtures that could reasonably be expected to occur in a giant planet, and see if we can determine the relative amounts of these substances. But how will we do this?

The answer is obtained by looking again at the process of planet formation. Recall from Chapter 4 that we think that the gases in substantial parts of the solar nebula were heated to a temperature well above 1000 K and then cooled, condensing a variety of solid materials in a series of steps as the nebular temperature decreased. As the temperature decreased from 1500 K down to 500 K, the main solids that formed were metallic iron, high-temperature silicate minerals such as Mg_2SiO_4, and sulphur compounds, of which the main one is FeS. As the temperature dropped further through the range of 500 K to 200 K, iron oxidized and was incorporated into silicates, and the minerals took on water and carbon to form rocks such as are found in the carbonaceous chondrites. Below 200 K, water ice began to form, and at still lower temperature, ammonia, methane, nitrogen, and carbon monoxide were able to form minerals in which these substances are embedded in a matrix of water (e.g. the hydrate $NH_3 \cdot H_2O$ and such clathrates as $CH_4 \cdot 6H_2O$, $N_2 \cdot 6H_2O$, and $CO \cdot 6H_2O$. The abundant elements H and He, however, were not condensed at any temperature that the solar nebula could reach.

This condensation sequence suggests a very simple division of chemical elements into three major generic substances that we should be able to treat as the main chemical components in an approximate description of the chemical composition of the giant planets. Let us call "rock" a mix of all the elements that condense above 500 K; this would include approximately solar proportions (see Table 2.2) of Mg, Si, Fe, and perhaps S, and enough O to fully oxidize the Mg and the Si. It is a little uncertain how much O we should include to oxidize iron; we certainly find both oxidized and metallic iron in various settings in the solar system. Because we are interested in the outer solar system, we should probably consider the iron to be oxidized. "Rock" would also include all the minor elements that condense at higher temperature, such as Ca, Al, Na, K, etc, and the O that would oxidize them (but these elements do not contribute much to the bulk of any planet).

A second component would be "ice", a mix of all remaining elements that condense above, say, 50 K, including in particular all the C (probably mainly as CH_4), the N (mainly as NH_3, and all the remaining O (mainly as H_2O). There would also be some N_2, CO, and CO_2, and a little Ne and Ar trapped in the icy matrix. The main feature of "ice" is that it probably contains all the O that did not condense with the "rock", together with solar proportions of C and N, and a tiny fraction of the total H.

The third component, which we call simply "gas", would be the gases that do not condense readily anywhere in the proto-solar system, essentially the He and the roughly 99.5% of the H that did not condense in the "ice" in combination with C, N, and O.

Note that we are using the labels "rock", "ice", and "gas" only to denote particular groups of chemical substances. Do not assume that each substance is in the physical state (solid, liquid, gas) suggested by its name! "Ice", for example, might be present as a solid in the interior of Saturn's moon Rhea, as a liquid under the crust of Jupiter's moon Europa (and in the Earth's oceans), or as a gas in the atmosphere of Jupiter.

Now it seems reasonable first approximation to consider constructing planet models in which we vary the proportions only of rock, ice, and gas. In this way we have only three materials for which we need to adjust the relative abundances, rather than dozens, so the problem of guessing a suitable composition for a planet is much easier. Furthermore, because we know that these three components were naturally produced, and frequently separated from one another, in the early solar nebula, we are using physically reasonable rather than arbitrary materials to build our theoretical planets. In fact, we have already effectively used one of these materials, rock, in our efforts to construct a plausible model of the Earth, and we found that the overall composition of the Earth is not too far from that of the rock that we are considering here. So now we examine more generally what kind of planets we can make from various proportions of these three materials.

Solving the general problem of the structure of a planet made from arbitrary proportions of our three basic materials requires considerable further information, about the behaviour of matter at high pressures. In a planet as massive as Jupiter or Saturn, the great weight of overlying layers will compress the material of the deep interior very strongly, much more than occurs inside the Earth. We need to know about the behaviour of our gas, ice, and rock mixtures at extremely high pressure. Since we will find that the outer planets are quite warm inside, we also need to know about the behaviour of these substances at fairly high temperature.

It is easy to get some idea of what an incredibly high pressure can be reached inside Jupiter. The pressure deep inside the planet must be great enough to support the weight of the material above each square meter. This weight, very roughly, is the product of the average mass of a single cubic meter (the mean density, about 1300 kg m^{-3}), times the height of the column of mass (the radius of the planet, about 70,000,000 m), times the planet's gravity (which at the surface is about 2.5 times greater than that of the Earth, about 25 m s^{-2}). The product of these three numbers is roughly 2×10^{12} Pa, or about 2×10^7 atm, more than 10 million times greater than the typical pressure in Jupiter's (or the Earth's) atmosphere. And as we shall see below, the temperature may be as high as 40,000 K inside Jupiter. So we are going to need information about the behaviour of H (or a mixture of H and He), for example, far outside the range of pressure and temperature covered by Figure 2.6.

We may also estimate the internal temperature inside Jupiter now that we have some idea of the pressure range. At the surface, the atmospheric temperature is measured to be about 170 K. It is found that if heat from the interior is carried outward by radiation or conduction, the excess heat that Jupiter radiates cannot be explained unless the temperature rises very rapidly inward, as both these mechanisms are rather ineffective at carrying heat inside Jupiter. But if the temperature rises rapidly inward, the planetary interior will become unstable to convection, as described in Chapter 10. Convection is a very efficient heat transport mechanism, and is entirely capable of carrying the excess heat out to the surface of the planet. Convection limits the rate at which the temperature can rise inward in the planet; if the temperature rises even a little more rapidly than is needed to keep the convection operating, so much heat is carried out that the temperature quickly adjusts to just the value that will continue the convection. Assuming that the planet is convective inside allows us to estimate that the central temperature of Jupiter will be about 4×10^4 K.

It is actually rather difficult to get the required information about the behaviour of gas, ice, and rock at very large pressures, especially when this is needed for a rather large range of temperature as well. Experiments are possible, for example by enclosing a miniature sample of a substance in between tiny diamond anvils which are then gradually tightened, or even by causing a small projectile to crash into a sample at extremely high speed, and studying the effects of the shock wave that travels through the sample. However, static experiments can only reach pressures of the order of 3×10^9 Pa. Impact experiments can get higher, up to 10^{11} Pa, but even this pressure is substantially less than the expected central pressure of Jupiter. The only possible means of studying the highest pressures at present is theoretically. For simple atoms such as H, this can be done reasonably exactly, but for more complex atoms and molecules such as are found in ice and rock, theoretical results are rather uncertain.

Two specific kinds of information are needed about our substances. First, we need to know about important phase changes that may occur as the pressure or temperature rises; recall the phase diagrams of Chapter 2, which describe water and H_2 (Figures 2.5 and 2.6). Secondly, we need to know how the density of each substance changes with increasing pressure and temperature.

For H, the main phase change of interest to us in connection with giant planet interiors is a transition which occurs at about 10^{11} Pa, where solid or liquid H_2 is so strongly compressed that the atoms find it energetically preferable to dissociate not only from each other (the molecules dissolve), but even to lose touch with their electrons, which join a general swarm of electrons spread out fairly uniformly between the H nuclei. This transition converts molecular H_2 into a structure which resembles a solid or liquid metal. The electrons, no longer attached to individual protons, but free to move throughout the material, conduct electricity and heat very well. We call this state **metallic hydrogen** .

A phase diagram for high-pressure, high-temperature H is shown in Figure 11.5. Compare this figure to Figure 2.6; note that Figure 11.5 has logarithmic (powers of ten) scales for T as well as P, to cover the wide range of values needed, and recall that a pressure of one atm in Fig 2.6 corresponds to $\log P = 5$ on Fig. 11.5. Figure 11.5 contains the boundaries between solid and liquid, and between liquid and gas, that are found in Figure 2.6, but these are now confined to a small corner of the figure. In a phase diagram covering a larger range of P and T, the solid-liquid boundary is seen to occur at constant T only for low pressure; once the pressure has risen above about 10^8 Pa, the boundary moves towards higher T. This reflects the fact that with increasing P, a solid is increasingly difficult to melt; the same effect makes the inner core of the Earth solid even though the cooler outer core is liquid. The boundary between solid molecular and metallic H is the horizontal line above 10^{11} Pa, but for a temperature above about 10^3 K the metallic phase is liquid. Two diagonal dashed lines in the figure near $T \sim 10^4$ K show the boundaries where rising temperature causes molecular H_2 to make the transition to atomic H, and then, at a slightly higher T, the H atoms become ionized.

The pair of dotted lines running diagonally through the figure show the approximate pressure and temperature combinations found inside Jupiter and Saturn, on the assumption that the internal temperature variation is determined by the occurrence of convection. Because of the limiting effect of convection on the rate at which T rises with depth, these lines correspond essentially to the highest T that could occur inside the two giants at each P. As noted above, the excess heat output of the two large giants over the input from the Sun leads us to believe that these “adiabats” actually describe the internal temperature of Jupiter and Saturn rather well. Examining the location of these two lines in the phase diagram, it appears that that the two planets have fluid molecular H_2 throughout their outer regions, but this gives way to liquid metallic H at great depth. These planetary adiabats miss by a wide margin the gas-to-liquid transition made by H_2 at low P and T, and shown in Figure 2.6; this indicates that as one descends into a H-rich atmosphere, there is no distinct

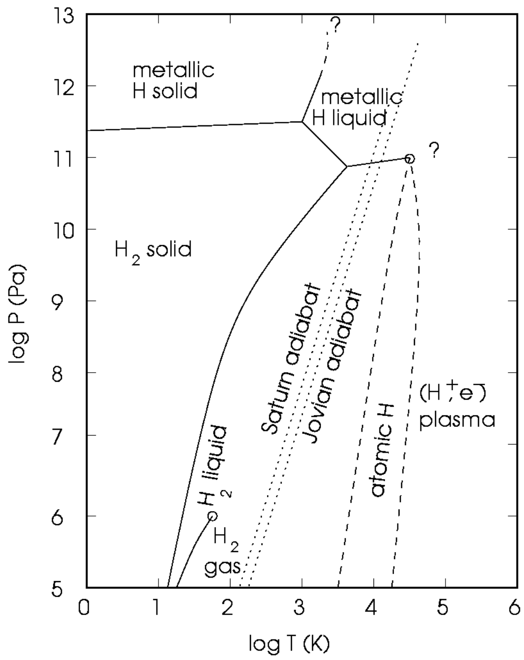

Figure 11.5: Phase diagram of hydrogen, for a much larger range of conditions than are shown in Figure 2.6. The solid and dashed lines show the phase transitions discussed in the text. The dotted lines indicate the range of conditions inside Jupiter and Saturn on the assumption that heat is carried outward by convection. $\log P = 5$ corresponds to 1 atm of pressure. (Adapted from Stevenson 1982, Ann. Rev. Earth Planet Sci., 10, 257.)

boundary where the liquid phase begins. The fluid simply becomes denser and denser as one descends. The atmosphere is "bottomless".

Similar combinations of experiment at (relatively!) low pressure and theory at higher pressure may be used to get some idea of the behaviour of the most water ice and other ices, and of various typical rock minerals. A further complication is that none of the chemical elements is likely to appear alone. H will be mixed with a lot of He (and some Ne, Ar, etc) in the gas component; water ice will be combined with CH_4, NH_3, CO, etc in the ice component, and the rock component will probably contain several minerals including olivine or pyroxene, and metallic Fe. This situation complicates the phase diagrams further.

The other kind of information needed, besides phase diagrams, is the variation of density with pressure for our various components. Such information is obtained from the same kind of combinations of experiment and theory as is used to define the phase diagrams. The way the density varies with increasing pressure for gas, ice and rock is shown in Figure 11.6. The curve for gas shows the variation of density for combinations of pressure and temperature estimated for the interior of Jupiter (see Figure 11.5). The density of the gas increases steadily with increasing pressure, even though at 10^9 Pa the molecules are certainly already in contact (i.e. the gas is really a liquid). At such high pressure, the molecules themselves are seriously deformed by the pressure. The densities of ice and rock are given in this diagram for slightly lower temperature at each pressure than is found in Jupiter; the conditions assumed are similar to those inside Uranus. However, neither the rock nor the ice density is very sensitive to temperature at these pressures, and the relationship shown would not be much different for the conditions inside Jupiter. Like the H-He mix, the ice component increases steadily in density as the applied pressure increases. The more resistant rock only begins to increase significantly in density above 10^{10} Pa.

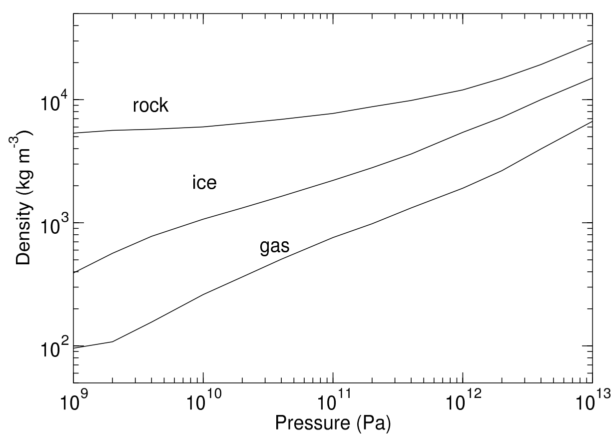

Figure 11.6: The relationship between pressure and density is shown for a solar H-He gas mix, for ice, and for rock. For the gas case, the temperature (not shown) increases with pressure approximately as shown in Figure 11.5 for Jupiter. For the ice and rock mixtures, the assumed T at each P is slightly lower than the curve for Saturn shown in Figure 11.5, and corresponds approximately to conditions inside Uranus. The density values graphed thus include significant effects due to temperature as well as increasing pressure. This is the reason that the density of ice is below 917 kg m^{-3}, its normal terrestrial value, over part of the graph. (Adapted from Stevenson 1982, Ann. Rev. Earth Planet Sci., 10, 257.)

Planetary models

Constructing models of the interiors of giant planets is actually easier than obtaining the information needed about the behaviour of the materials that go into the planets. We must solve a set of equations like those used in getting models of the terrestrial planets. These equations describe how the attraction of gravity at each level varies, depending on the mass inside that level; how the pressure increases, due to the weight of overlying layers; how the temperature varies, due to the effect of convection which carries heat out; and of course, the way in which the density of the assumed material increases with increasing pressure and temperature.

The simplest situation for which to construct models is for material at $T = 0$ K. Figure 11.7 shows how the radius of the resulting models depend on the assumed mass of the planet for planets made of pure gas, pure ice, and pure rock. For a planet made of H or an H-He mixture, the radius of the planet increases rapidly with mass up to about 300 Earth masses, and then stops increasing, as the effect of the large mass acting to compress the gas becomes really important. For larger masses, the size actually starts to decline with increasing mass; this behaviour is characteristic of white dwarf stars, which have the unexpected behaviour that the larger the mass of the star, the smaller it is. Cold planets of ice also increase in size rapidly at first and then more slowly, as increasing mass compresses the material strongly. Rock planets, more resistant to compression, increase steadily in size with increasing mass throughout the mass range shown.

Notice that the low density of H relative to other substances leads to a gas planet being *much larger* at any particular mass than a planet of rock or ice. Each of the four giant planets is plotted at the appropriate point for its mass and size (points noted as J, S, U, and N), and it is clear that if Jupiter and Saturn can be treated approximately as cold planets, they must have compositions dominated by H and He. Even a pure ice planet is far more compressed than either of these planets. On the other hand, Uranus and Neptune could plausibly be modeled as planets made largely of ice (or ice and rock), with only a modest external layer of gas.

Models with more realistic temperatures have also been computed, again for pure gas, ice, and rock. Using the pressure-density relationship of Figure 11.6, one obtains the models shown with dashed lines in Figure 11.7. The size of the rock models is hardly increased at all by the effect of temperature. The ice models are increased in size by a few percent. The gas models, however, are drastically changed, at least for low mass, where the "fluffiness" of H_2 gas makes a warm model much larger than a cold one of the same mass. As the mass of the gas model increases, this effect diminishes as the increasing gravity of the planet compresses the hydrogen. At a mass of 300 Earth masses, appropriate to Jupiter, the radius increase caused by finite temperature is only about 10%. It is still clear that among the substances we have chosen to discuss, Jupiter and Saturn must be composed primarily of gas. The fact that both lie a little way below the line for "warm" models of the gas (H-He) mixture means that neither planet can be pure H-He; each must have a modest amount of denser material (some combination of ice and rock) somewhere inside.

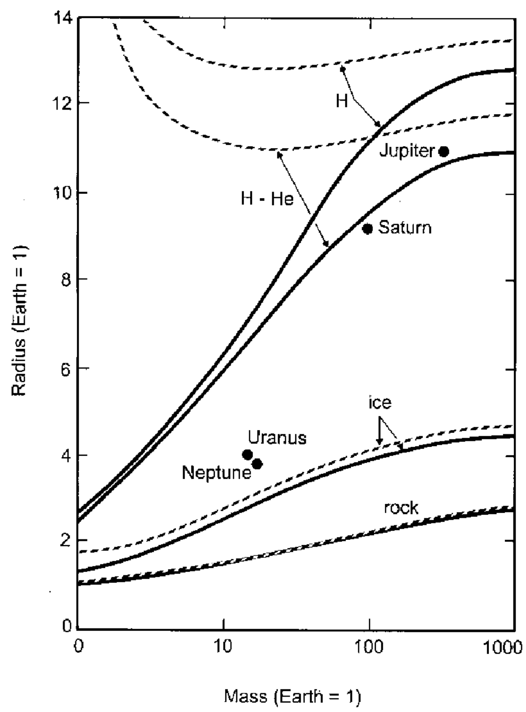

Figure 11.7: The relationship between mass and radius for model planets. The solid lines describe planets at $T = 0$, while the dashed lines correspond to models calculated with the density-pressure relationships of Figure 11.6. The positions of the gas giants in this diagram are indicated by their names. Notice that the mass scale is really logarithmic, while the radius scale is linear. (Adapted from Stevenson 1982, Ann. Rev. Earth Planet Sci., 10, 257.)

Exercise: Why is the difference between "warm" and "cold" models *not* an important distinction for terrestrial planets?

The next step in modeling is to consider layered models, in which for example a shell of gas surrounds a core of rock and ice. In this case the mass of the rock and ice in the core must be determined. One may also choose to vary the ratio of He to H in the gas envelope, to try to determine this quantity directly rather than assuming it. With more free parameters, these models cannot be easily shown in a single graph like Figure 11.7. However, such models begin to approach the level of complexity needed to agree with all the constraints available from observation (mass, radius, the degree of flattening of the planetary shape produced by the rapid planetary rotation, and the external gravity field as measured by space probe fly-bys). All four giant planets have been described by such layered models. Remarkably, all four are found to have dense cores of rock or ice with masses of the order of 10 to 20 Earth masses. They vary in mass and size primarily in the mass and extent of their H-He envelopes, which range from about 20% of the total mass for Uranus and Neptune up to 80% for Saturn and 95% for Jupiter. Both Jupiter and Saturn have molecular H_2 in their outer layers and metallic H inside that, but in Jupiter the metallic H makes up a much larger fraction of the H-rich envelope than in Saturn. In Uranus and Neptune, with less massive gas envelopes and smaller total masses (and therefore internal pressures), H is everywhere molecular H_2.

It is interesting to look at these models more closely. The models for Jupiter all agree on the general structure of the largest planet, although they vary significantly concerning the details. An overview of a typical Jupiter model is shown in Figure 11.8. With it, we may imagine a guided tour of this strange world.... We enter Jupiter from above the atmosphere, through a gas of predominantly H and He in which both density and temperature rise rapidly with depth. Where the pressure is comparable to that at ground level in the Earth's atmosphere, we encounter the first cloud deck, the level at which rising and cooling convection currents reach a low enough temperature for NH_3 to freeze out as ice crystals. Depending on where we enter the atmosphere, the clouds may be yellow, tan, reddish, or white. As we go deeper into the atmosphere, we come to a second cloud deck of ammonium hydrosulphide (NH_4SH) ice, and then a third cloud deck, of water ice. As we continue inwards, the pressure, density, and temperature all rise. By the time we are 700 km below the highest clouds (0.1% of the radius of Jupiter), the temperature has risen to 2000 K. No gas-liquid interface is encountered. However, the H/He-rich atmosphere becomes steadily denser, and at the level where $T \approx 3000$ K and $\rho \approx 100$ kg m^{-1} the fluid in which we are descending resembles a liquid more than a gas. Here the molecules touch one another, and compression of the fluid is much more difficult than it was high in the planet's atmosphere. The fluid is slowly convecting, carrying heat towards the surface.

Slightly below $0.8R_{\rm J}$, about 140,000 km down from the clouds, we reach a point where $\rho \approx 1000$ kg m^{-1} and $T \approx 10^4$ K. At this level the pressure and temperature are high enough to **dissociate** H_2 molecules into single (atomic) H atoms. This is the first major transition zone; the density rises significantly, though perhaps not abruptly. We continue downward with little change except for steadily rising temperature, pressure, and density, until we reach a depth of about 54,000 km ($0.2R_{\rm J}$ from the centre) where the surface of the rock core is finally encountered. The temperature here is about 20,000 K. The mass of the rocky core is somewhat uncertain, and it is not know how much ice it includes, but it very likely contains between 10 and 30 $M_\oplus$ (Earth masses). Note that even if we separated all the chemical elements heavier than He out of the H/He envelope (and the presence of CH_4, NH_3, etc in the atmosphere shows that "ices" at least are partly dissolved in the gas envelope), we would be able to make a core of rock and ice of at most about 5 $M_\oplus$. Jupiter clearly has *excess* rock over a purely solar chemical composition. However, only about 5% of the planet resembles a terrestrial planet; the other 95% is in the huge H/He atmosphere.

The structure of Saturn is similar to that of Jupiter, except that observations suggest that the H/He gas mixture has only about one He atom for every 20 H atoms instead of the usual (solar) ratio of about 1 in 10, which is found in Jupiter's atmosphere. We will return to this observation below. As we descend into the atmosphere of Saturn for a tour, we encounter the same series of cloud decks as on Jupiter: a first layer (pale yellow in colour) of ammonia ice crystals, a second of ammonium hydrosulphide ice, and a third of water ice crystals. As we descend into the planet, the density, pressure, and temperature rise steadily. As in the atmosphere of Jupiter, the temperature increase with depth is rapid enough to cause convection, which carries the minor chemical species NH_3, etc to the altitude where the temperature is low enough for the molecules to condense as ice crystals and fall back down. Again as we descend, no abrupt change of phase is encountered; there is no sea below the atmosphere. The atmosphere into which we descend simply becomes more and more dense (and hot). Where the density reaches 50 – 100 kg m^{-3}, the H_2 molecules are in nearly continuous contact with one another, and the fluid makes a seamless transition to a liquid state. At about $0.5R_{\rm J}$ from the centre, the molecular liquid is transformed by increasing pressure and temperature into an ionic gas, as the molecules of H_2 dissociate into free protons immersed in a swarm of detached electrons. It is thought that in this region He condenses into "raindrops" which fall through the H gas. It is this separation and settling of He which is thought to have gradually depleted the stock of He in the outer envelope of Saturn. At about $0.25R_{\rm S}$ from the centre, where the temperature is about 14,000 K, we finally reach the predominantly rock core

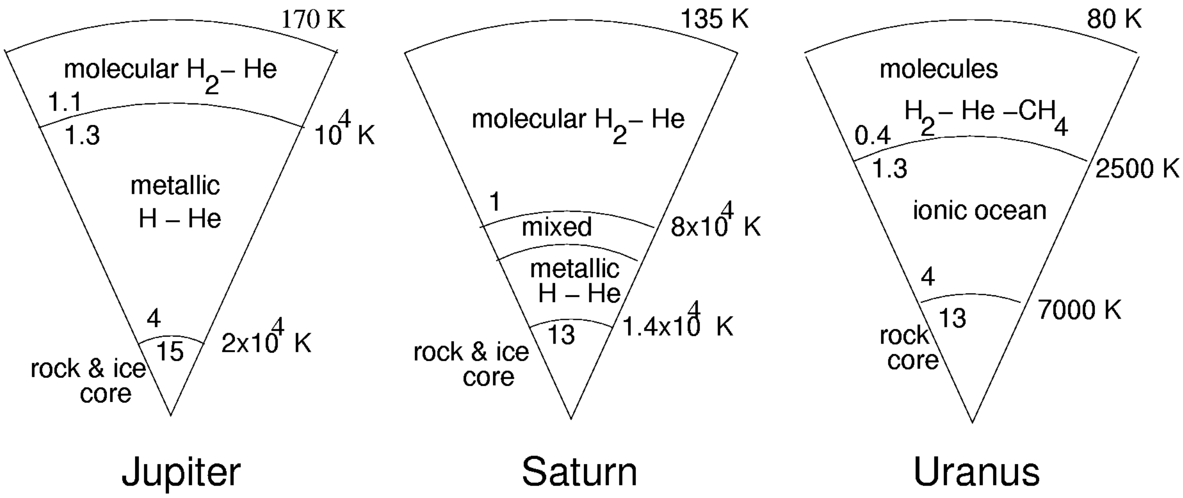

Figure 11.8: Sketches of the interior structure of Jupiter, Saturn, and Uranus. Boundaries between regions are shown at the appropriate fraction of the radius of each planet from the centre; recall that the planets actually have different sizes! Numbers adjacent to boundaries near the left side of each figure are approximate densities, in gm cm^{-3}; those near the right side are approximate temperatures at the boundaries. (Adapted from Stevenson 1982, Ann. Rev. Earth Planet Sci., 10, 257.)

of the planet. As in Jupiter, the total mass of this core is uncertain, but it is believed to be about 20 $M_\oplus$, far more than the 1 – 2 $M_\oplus$ that could be obtained by separating all the heavy elements from a mix of Saturn's mass and solar chemistry.

Uranus and Neptune are similar to one another. They differ from Jupiter and Saturn primarily in having rock cores of much greater relative importance. The rock core of each planet is probably surrounded by an ocean of liquid ice, above which is found the gas atmosphere. The blue colour of these two atmospheres are due to the gaseous methane mixed in with the dominant H and He molecules; this gas strongly absorbs red sunlight, leaving mainly bluish light to be reflected back to the outside observer. As we descend into the atmosphere of one of these planets, (Uranus, say), we first encounter the cloud deck of methane (which cannot freeze out to make clouds anywhere in the atmospheres of Jupiter or Saturn because the temperature is too high at all altitudes). The other cloud decks found in the two largest giant planets occur in the same order below this first cloud layer. Again we descend through an atmosphere which for thousands of kilometers is bottomless, in which temperature and density steadily rise. At about $0.7R_U$ from the centre, where $T \approx 2500$ K, we encounter the surface of an immense ionic ocean of $H_3O^+OH^-$ (ionized water) with dissolved NH_3. The surface of this ocean is marked, not by a gas to liquid transition, but by the abrupt change of chemical composition, and by a change of density from about 400 to 1300 kg m^{-3}. The floor of this ocean, which is some 10,000 km deep, occurs at about $0.3R_U$, where $T \approx 7000$ K. This is the beginning of the rock core of the planet, an object of a few Earth masses. It is possible that the core is differentiated into a silicate mantle and an iron core as in the interior of the Earth.

The solar nebula and the origin of the giants

We now turn to the problem of understanding the formation of the giant planets. It is generally believed that formation of these planets took place during the period when the Sun was forming, out of the material that was orbiting the Sun in the solar nebula. Most of the matter that at one time was in the solar nebula ended up in the Sun, of course. The planets formed from some of the matter that did not, as discussed in Chapter 4.

The material that was to become the Sun and the solar system separated from the parent giant molecular cloud by one or more episodes of gravitational collapse. Some of this collapsing gas, with its embedded dust, fell more or less straight in to the centre of the shrinking cloud fragment, and quickly became part of the forming Sun at the centre. However, most of the gas was rotating too rapidly about the centre of the cloud fragment to fall directly into the centre. Instead, this gas and dust settled into a thick disk-shaped cloud, tens of AU in radius, around the growing central Sun. This disk was the solar nebula.

The disk of gas around the Sun was fairly **turbulent**:

on a large scale, the gas was moving in roughly circular orbit around the Sun, but various individual blobs of gas had somewhat eccentric motions above and below the average plane of the nebula, and in and out. This turbulent, disorderly state meant that individual gas blobs collided with one another frequently. The collisions tended to alter the motion of a gas blob to be more like the motions of other blobs at the same distance from the Sun, and thus these collisions tended to diminish the turbulence by encouraging all the gas to orbit in smooth circular orbits. However, a considerable degree of turbulence was maintained in the disk by blobs and cloud fragments falling into the solar nebula from the parent giant molecular cloud as the Sun developed.

The solar nebula was much warmer than the giant molecular cloud from which the cloud fragments collapsed. This was mainly due to the energy released by fresh gas blobs falling into the nebula, and by the friction of cloud fragments colliding with one another. As the central Sun became more massive and luminous, its radiation also heated the nebula. There is evidence from the meteorites (remember the Ca-Al rich white inclusions in some carbonaceous chondrites?) that some parts of the nebula reached temperatures of 2000 K or so. As the Sun formed, the temperature in the nebula varied from place to place (and also with time, of course), from a maximum of perhaps 2000 K near the Sun, diminishing to perhaps 50 K at the outer edge of the nebula. In the inner parts of the nebula, rock first vaporized and then (as the nebula cooled) condensed as dust grains, while both ice and gas remained gaseous until the process of planet formation was finished. In contrast, in the outer part of the nebula, the generally lower temperature allowed both rock and ice to condense into solid grains.

The turbulence of the nebula insured that tiny grains of condensed solids collided frequently with one another. They must have often stuck together. The size of typical grains grew more or less steadily because of these collisions, gradually producing larger and larger objects. Eventually, we imagine that the solid bodies would have ranged in size from smaller than a grain of salt to larger than an iceberg.

Now we can imagine two rather different ways in which the giant planets could have formed from this rotating, turbulent solar nebula and the many orbiting solid objects it contained. One possibility is that the growth of solid bodies by collisions could have gone on until planetesimals of rock and ice several times more massive than the Earth were formed in the region of 5 to 30 AU from the proto-Sun. Once the largest planetesimals became this large, they would have such powerful gravitational attraction form nearby material that they would begin to accrete *uncondensed* H- and He-rich gas, eventually growing to the mass of the present giant planets. A second possibility is that within the solar nebula, conditions allowed large cloud fragments to collapse directly due to their internal gravitational self-attraction. That is, one might have had within the solar nebula the same kind of gravitational fragmentation and collapse that occurred on a large scale in the giant molecular cloud from which the Sun formed.

The creation of the giant planets by direct gravitational collapse now seems much less likely than the formation first of cores of rock and ice by collisions, followed by sweeping up of gas envelopes from the nebula. The main problems with the direct collapse theory are the following. First, the direct gravitational collapse hypothesis does not offer any obvious reason for the existence of rock (and ice?) cores in the giants; if these planets formed by direct collapse, heavy elements would have tended to remain dissolved in the hot convecting gaseous envelope rather than settling to the centre. Furthermore, the direct collapse theory offers no explanation for the fact that the four giant planets are so different in *total* mass but nevertheless all have similar *core* masses. A related difficulty is that the giants all have excess rock and ice compared to gas, with a *larger* excess of rock in the outer giants. This requires either extraordinary accretion of solids after the collapse phase, or the loss of much of the gas afterwards, with the fractional gain of solids, or the loss of gas, *increasing* towards the outer edge of the solar system. The collapse theory does not provide any explanation for an increasing departure from solar chemistry farther out in the solar system.

Thus, we consider the idea that in the outer solar system, solid planetary cores formed by collisions, and that once they reached some critical size they began to rapidly sweep up the gas and dust around them. This would have increased their gravitational attraction, increasing their sweeping efficiency, until each planet effectively cleared out a large band around its orbit. Computations simulating this process indicate that a core mass of the order of 10 $M_\oplus$, about the size of the cores actually present in the giant planets, would be the mass at which rapid accretion of nebular gases would begin to occur. This process *does* account for the similar core masses in all four giants. Furthermore, since the cores form before gas accretion begins, they are naturally segregated in the planetary centres. However, it is not yet clear what limited the masses to which each giant planet grew. Perhaps accretion of nebular gas came to a halt as the remaining gas was ejected from the solar system by a T Tauri-like stellar wind. Or perhaps the masses of the giants are limited by the total mass of gas in rings swept by each core in the outer solar system at the time when the cores reach a mass large enough to begin "vacuuming up" the nebular gas. In any case, the decrease in density in the nebula with increasing distance for the Sun makes it likely that the core of the inner giants probably accumulated more quickly than those of Uranus and Neptune, and in a denser region of the solar nebula, so that the much larger gas envelope masses of Jupiter and Saturn compared to the two outer giants seems natural. One remaining puzzle, if the outer planets formed from massive "rocky" cores, is why the cores were so much more massive in the outer solar system than in the inner, where the most massive planets only reach 1 $M_\oplus$.

Evolution of the giant planets

We have seen that the giant planets probably originated from rocky proto-planetary cores that grew to masses of roughly 10 $M_\oplus$ through successive collisions of small solid bodies in the outer solar nebula. Let's look now in more detail at how these cores accreted their massive envelopes, and how the structure of these developing planets changed with time.

Early in the process of planet formation, the solar nebula was swarming with small rocky planetesimals. As these bodies collided with one another and coalesced, the dimensions of the largest objects grew as their number declined. The largest bodies gradually increased in size and mass, first to asteroidal dimensions, then to the sizes of terrestrial planets, and finally, in the outer solar system, to masses of about 10 $M_\oplus$. As these largest rocky bodies were growing in the outer solar system, they were also beginning to sweep up and retain gaseous atmospheres from the gas of the solar nebula. Because the rocky cores were moving within the nebula, there was no sharp outer limit to their growing atmospheres. Instead, these atmospheres merged smoothly with the nebula. The outer limit of the atmosphere of one of these large proto-planets simply occurred at a distance from the planet at which the gas ceased to be gravitationally attached to the rocky core, and was instead controlled by the tidal attraction of the growing Sun. In effect, the Sun's gravity forced gas at large distances from the proto-planet to orbit the Sun at a different speed than the proto-planet, rather than remaining close to the planetary core. However, at such large distances from the Sun (between 5 and 50 AU), the size of the region gravitationally controlled by the planetary core was enormous, and so the radius of the thin planetary atmosphere, and the effective size of the growing planet, was hundreds of times larger than the present radius of Jupiter.

As the largest cores grew in mass, the amount of gas that a core could attract and hold also grew steadily. The gas atmosphere grew hotter as it grew in mass, due to the gravitational energy released as material was drawn steadily closer to the growing core. An approximate equilibrium was established, with the gas pressure in the atmosphere at each level (a pressure produced by the compression and heating of the gas) having about the right value to support the weight of overlying gas layers. As more and more gas became attached to the atmosphere inside the growing sphere of influence of the core (which continued to grow in mass by collisions), gravitational compaction of the atmosphere steadily raised its density and temperature so that the pressure balance continued to be maintained.

However, once the internal temperature of the gas atmosphere rose to about 2500 K, at a time when a few Earth masses of gas had been accreted, the H_2 molecules deep in the atmosphere began to *dissociate* into H atoms. This process absorbed a lot of heat, and prevented the temperature of the gas from rising significantly further until the H_2 was mostly dissociated. This in turn upset the pressure balance supporting the huge H/He atmosphere, and this atmosphere quickly shrank to a much more compact state. The shrinkage in turn reduced the pressure holding away the surrounding nebular gas, and much of the nebular gas flowed into the gravitational sphere of influence of the protoplanet. As fresh gas filled the planet's sphere of influence, the total planetary mass increased rapidly, further enhancing the ability of the protoplanet to attract still more nebular gas. The result of this collapse was the rapid accretion of anywhere from about 1 $M_\oplus$ (on Uranus and Neptune) to tens or hundreds of Earth masses (Jupiter and Saturn) of atmospheric gas. It is this stage of the accretion process that can be thought of as a kind of "vacuuming up" of the available nebular gas by the rapidly growing planet. At the end of this process, all of the giant planets had reached roughly their present masses.

At this point, the giant planets were in other ways rather different from their present structures. All four giants were considerably hotter inside than they are today, due to the large amount of gravitational energy released by the accretion of the massive gas atmospheres. As a result, they were several times larger in diameter than they are today, and radiated far more infrared energy than now. Jupiter, the most extreme case, probably radiated about ten million times more energy per second than at present; its initial luminosity was about 1% of the Sun's luminosity!

In the 4.5×10^9 years since they were formed, the masses of the giant planets have remained essentially constant. Their structures, rocky cores surrounded by huge gaseous atmospheres, have not changed qualitatively. However, each planet has reached a sufficiently compact state that as heat leaks out from the interior, it is not *not* replenished by further shrinking and release of gravitational energy. Instead, the heat flowing out from the interior and radiating away from the surface simply lowers the internal temperature of the planet. As each giant cools, it does shrink slightly, and as the internal heat energy supply decreases, the intrinsic luminosity (the excess of the planet's luminosity over the input of absorbed solar radiation) slowly decreases also.

When the cooling history of each of the outer planets is calculated, it is found that Jupiter, Uranus, and Neptune all have about the expected radii and luminosities. This fact supplies a powerful confirmation of our ideas about the internal structures of these planets, and our picture of how they have evolved during the 4.5×10^9 age of the solar system.

However, Saturn is about twice as luminous as it is expected to be. It appears that a second important energy source is present inside the planet in addition to the energy available as internal heat released by the formation and subsequent contraction of the planet. This extra energy is probably released by the separation and settling of the He relative to the dominant H of the gas envelope. What we think happens is that as the temperature in the outer layers of metallic H in Saturn drops below about 10^4 K, He becomes partly *immiscible* in (unable to remain dissolved in) the H. Droplets of He form, and since they are denser than the surrounding fluid H, they fall towards the centre of the planet as a kind of exotic hot rain. This separation is thought to release enough energy to account for the part of Saturn's current infrared radiation that is not due simply to solar heating and global cooling of the hot planet. The separation of He from H in the deep interior of the Saturn gradually leads to a general depletion of He throughout the outer envelope, and is probably responsible for the observation (see Sec. 10.2) of a deficiency of He in Saturn's atmosphere compared to the ratio found in the Sun and in the other giant planets.

The interesting fact that Uranus seems not to be radiating any more energy than it absorbs from the Sun each second, while the more distant and quite similar planet Neptune is still losing internal energy, is simply due to the larger input of solar heat into Uranus compared to Neptune. The solar heat input now holds the surface temperature of Uranus approximately constant at the observed 57 K. As the interior continues to cool, the radiating surface can no longer cool at the same pace, and so the efficiency of internal heat transport is greatly reduced, leading to a large decrease in the transport of internal heat to the surface. As a result, internal heat loss now makes up only a small fraction of the total radiation from Uranus.

Exercise: Could a giant planet with a mass of, say, 20 Earth masses have a density as large as that of the Earth? If so, how would such a planet have to differ from the structure of Uranus and Neptune?

11.3 Moons of the giant planets

Observations and exploration

The first moons discovered around a giant planet were the four largest moons of Jupiter, found by Galilei in 1609 with his small telescope (Chapter 1), and easily visible through ordinary binoculars. These "Galilean" moons, named Io, Europa, Ganymede, and Callisto after lovers or companions of Jupiter in classical mythology, are all comparable to the Earth's Moon, with radii ranging from 1560 to 2630 km (similar to the Moon's radius of 1740 km). Saturn has one moon of this size, Titan, found by Christiaan Huygens in 1655. The two largest moons of Uranus, Titania and Oberon, were first seen in 1789 by the planet's discoverer, William Herschel. Neptune also has one large moon, Triton, found by English astronomer William Lassell in 1846, not long after the discovery of the planet itself. These large moons are rather similar in many ways to the terrestrial planets; the largest, Ganymede and Titan, have larger diameters than Mercury, although they are only about half as massive.

During the past two centuries, many somewhat smaller moons having radii of 150 – 750 km were discovered by Earth-based astronomers, and at the start of the 1970's the number of known satellites of the giant planets was over 20 (Table A.3). This number was more than doubled with the discovery between 1979 and 1989 by the Voyager spacecraft of more than two dozen smaller moons, bodies with radii of the order of 50 km or less. It has gradually become clear that each of the giant planets has a *system* of moons, anywhere from eight to 18 or more moons orbiting the main planet. An important and striking feature of these moon systems is that in general the nearer and larger moons travel in nearly circular, concentric orbits around the equator of their planet. The great regularity of the orbits of the moons provides us with an important clue to their origins, as we shall see below.

Although numerous moons were discovered using Earth-based telescopes, little more could be learned about them in this way. The largest moons of Jupiter are not much more than a second of arc across as seen from the Earth, about the same as the typical blurring of images by the Earth's atmosphere. Even using special techniques developed in the past decade to reduce the blurring effect of the atmosphere, or using the

Hubble Space Telescope, one can at best make out the largest features on the nearest large moons. Because detailed images of even the largest and nearest moons of the giants could not be obtained from the ground, even the diameters of these moons were very poorly known before the first space missions to the outer planets. Nothing at all was known about surface features.

The two Voyagers and the Galileo space probe completely changed the situation. These spacecraft carried instruments for several kinds of measurements. Each carried a camera, and we now have an enormous collect of spectacular, detailed images of the surfaces of many of the moons of the giants. Each was also equipped with one or more spectrographs which were used to study both the chemistry of the atmospheres of the planets themselves, and the surfaces of the moons. Several instruments in each probe provide information about the particles and magnetic fields in interplanetary space and in the vicinity of the planets. Finally, the radio signals sent back to Earth by these probes are also studied to provide information about the precise positions and movements of the spacecraft as they pass close to the planets and the largest moons.

Approximate masses of a few of the larger moons had been determined before the era of space exploration by careful study of the apparent motions of various moons about their planets. The orbit of each moon is influenced (perturbed) somewhat by the gravitational pull of the largest nearby moons, and observation of this effect allows one to deduce the masses of the larger moons. More precise masses for a number of the moons were obtained from the Voyagers and Galileo by careful study of the changes in frequency (the Doppler shift) of the radio transmissions from the spacecraft as it passed close to a moon. These frequency changes enabled mission scientists to deduce the acceleration of the spacecraft by the gravitational pull of the moon, from which the moon's mass could be determined. In all, the Voyagers provided new mass measurements for 17 of the moons.

Another extremely important kind of information provided by the Voyagers and Galileo was accurate measurements of the diameters of all but the smallest moons, and determinations of the shapes of some of the smaller irregular objects. (Most of the smallest moons appeared to the Voyagers only as points of light even at closest approach, making accurate size measurements impossible.) The diameters that were determined allowed accurate determinations of the mean densities for about 20 of the larger moons. The mean density, of course, provides us with a very powerful clue to the bulk chemical composition of the object. It is found that the two inner Galilean moons of Jupiter, Io and Europa, both have densities of more than 3000 kg m^{-3}, similar to the density of rock (about 3500 kg m^{-3}). All the other large satellites have densities lying between 2100 and 1000 kg m^{-3}, low enough that these objects cannot be made solely of rock, but must have at least half of their mass made up of ice (density about 920 kg m^{-3}). Most of the large moons apparently are composed of 60 or 70% ice by mass. An important problem, and one which remains unsolved to a considerable extent, is to determine which of these large moons have differentiated into a structure having a rock core and an ice mantle, and which remain more or less uniformly mixed.

Because the moons of the giant planets are solid objects like the terrestrial planets, and generally lack significant atmospheres (although Saturn's largest moon, Titan, has a dense N_2-rich atmosphere, and Neptune's Triton has very tenuous atmosphere), the Voyagers and Galileo have returned a wealth of information about the surfaces of these bodies in the form of thousands of detailed photographs. The large moons were revealed to be an astonishingly diverse group of objects, many of which have histories comparable in complexity to the history of Mercury or the Earth's Moon. Most of the moons have surfaces which are more or less heavily cratered, but two (Io and Europa) are essentially totally free of cratering. Several other moons, in contrast, have a largest crater not a great deal smaller than the moon itself, while still others have few craters with diameters of more than about 50 km. Some moons have huge rift valleys, fractures, scarps, or flows. Two (Io and Triton) have active volcanos. A number of kinds of terrain are unfamiliar and do not yet have generally accepted explanations.

Jupiter's moons

Jupiter has four distinct systems of moons. Close to the planet, between 1.8 and 3.1 Jupiter radii ($R_{\rm J}$) and within the faint planetary rings, are four tiny satellites. All travel in direct, nearly circular orbits close to Jupiter's equatorial plane. The innermost two are even within the planet's Roche limit, so they must have some internal strength in order not to be disrupted by tides.

A second system is composed of the four large moons discovered by Galileo Galilei. These also move in nearly (but not exactly) circular orbits within one degree of the planet's equatorial plane, between $5.9R_{\rm J}$ and $26R_{\rm J}$ from Jupiter's centre. The orbital periods (recall that the period is the time taken for one revolution around the central planet) of the inner three moons show a remarkable kind of resonance: each time Ganymede makes one complete circuit around Jupiter, Europa makes almost exactly two trips, and Io makes four. The motions are synchronized in such a way that each time Io passes Europa (which happens once for each complete orbit of Europa), Io is at the point in its (slightly elliptical) orbit closest to Jupiter, while Europa is at the point in its orbit farthest from the planet. The moons thus jostle each other in a regular way, which is what maintains the slight eccentricity of the orbits.

The two outer systems of small moons include four at about $160R_{\mathrm{J}}$ in rather eccentric orbits, inclined to Jupiter's equator by nearly 30°, and four in seriously eccentric *retrograde* orbits at around $300R_{\mathrm{J}}$ – these outermost moons orbit the planet in the opposite sense from the inner moons and the planet's own rotation.

Little is known about the physical natures of the small moons. The sizes of the four inner moons are known – all are far from round, with dimensions in the range of about 10 to 120 km – but even this information is quite uncertain for the outer moons. The small moons are all dark and reddish in colour. Distant images from the Galileo orbiter show no surface features on any of the inner moons except for craters.

The four Galilean moons are much more varied and clearly have had complex histories. From the accurate radii derived from probe images, and masses determined from the gravitational deflection of the Voyagers and the Galileo probe by the individual moons, we now have very accurate mean densities for all four moons. Io, with a density of about 3500 kg m^{-3}, must be composed essentially of rocky material, probably similar in composition to the rock component we discussed in connection with the planets themselves; it may have a significant iron core as well. Europa has a density of about 3000 kg m^{-3}, and is thus primarily composed of rock and perhaps some iron, but ice probably makes up roughly 20% of the total mass of the moon. The other two moons have densities a little below 2000 kg m^{-3}, and probably are made up of about 40% rock and 60% ice. The density of the four moons falls systematically with distance from Jupiter.

From careful examination of the gravitational deflections of orbiters during close passes by a moon, it is also possible to get some information about the degree to which matter is concentrated towards the centre of a moon. This information helps to decide whether a moon is differentiated, with a core-mantle structure, or is homogeneous, with its materials uniformly mixed. (Note that some central condensation will occur even in a large homogeneous moon because of compression of matter near the centre by the weight of overlying layers.) Such data definitely indicate that Io, Europa, and Ganymede are differentiated. It is not yet completely clear if Callisto is "somewhat" differentiated or still homogeneous: although the bulk composition of Callisto is similar to that of Ganymede, Callisto clearly does *not* show the same degree of concentration of high-density matter to the centre that the larger moon exhibits.

Models of the three differentiated moons suggest that Io probably has a core of iron and iron sulphide (Fe-FeS) extending from the centre out to somewhere between $0.35R_{\mathrm{I}}$ to $0.60R_{\mathrm{I}}$ (R_{I} is the radius of Io), surrounded by a mantle of silicate rock extending to the surface. Europa appears to have a crust of water ice between 80 and 200 km thick surrounding a silicate rock mantle. It probably has an Fe-FeS core extending out from the centre to somewhere between $0.30R_{\mathrm{E}}$ and $0.50R_{\mathrm{E}}$. Ganymede, which has a lower mean density than the two inner Galilean moons, probably has a layer of water ice about 800 km thick over a rocky silicate mantle. It seems likely that this mantle overlies an Fe-FeS core having a radius of between $0.15R_{\mathrm{G}}$ and $0.50R_{\mathrm{G}}$. The presence of magnetic fields in Io and Ganymede support the view that these moons have metallic cores.

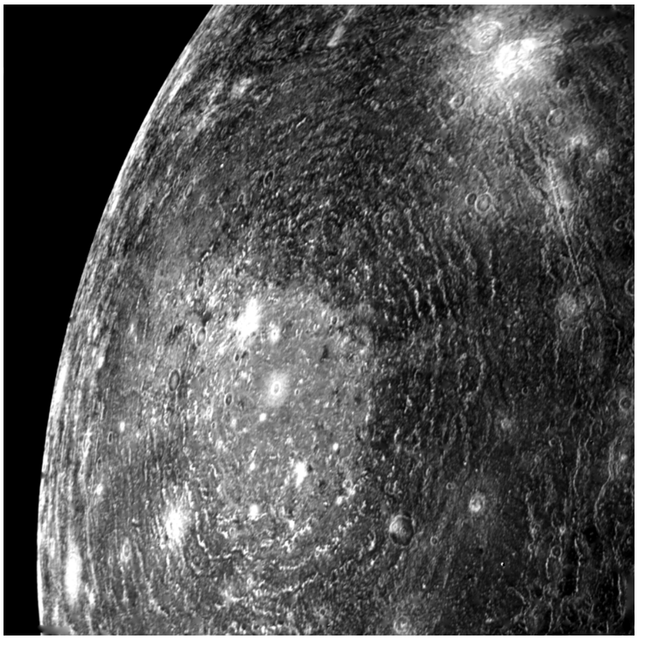

Figure 11.9: This image of Callisto acquired by the Voyager 1 space probe shows the huge impact basin Valhalla. This crater is surrounded by numerous rings that are mainly visible due to colour variations; they have very little vertical relief. This basin resembles the Orientale basin on the Moon, or Caloris on Mercury. The only other visible features are smaller craters. (Courtesy of NASA.)

The surfaces of these four moons are highly distinctive. Callisto, the outermost Galilean moon, and the least differentiated, is a heavily cratered body which shows no clear signs of any past volcanic or tectonic activity. The largest crater on the moon is about 600 km in diameter, about 1/8th of Callisto's diameter, and is surrounded by concentric rings, rather like the Orientale Basin on the Moon (Figure 11.9). Many of Callisto's craters are rather flat, as though the surface is not able to support high relief. This fact suggests that the surface layers may contain large amounts of easily melted water ice, consistent with a bulk composition of more than half water deduced from the mean density, and with spectra of reflected sunlight. Much of the surface is rather dark, so that surface water ice must be mixed to a considerable extent with some other material. The youngest craters are quite light in colour, which suggests that these impacts have excavated relatively clean ice from below the surface. Perhaps a thin surface layer was actually melted early in Callisto's history, allowing rocky material to settle below the surface, and subsequently contaminated with more rocky material by collisions with small asteroids and comets. A puzzling feature of the crater record on Callisto is that there are not as many small craters as one would expect from the number of large craters seen; instead, some smooth, dark material blankets the surface in between craters (Figure 11.10).

Figure 11.11: A full disk image of Ganymede taken by the Galileo orbiter during its first close encounter with the moon in 1996. Notice the division of the surface into two strongly contrasting terrain types, a dark terrain type which on close inspection is found to be heavily cratered, and a light-coloured surface which is observed to be much less heavily cratered but covered with a network of grooves and furrows. (Courtesy of NASA.)

Ganymede is similar in size and overall composition to Callisto, but unlike Callisto, the rocky and metallic materials have settled to the centre beneath the ices, leading to a core-mantle structure. The surface, which in some parts is heavily cratered like that of Callisto, also shows large regions with relatively light cratering that are intensely furrowed or grooved. In spite of similar size and bulk composition, the two moons have clearly had remarkably different histories. The surface of Ganymede, in fact, still presents us with serious difficulties in understanding how the features we see were created. There are two major terrain types present on Ganymede (Figure 11.11). Old regions of the moon, about 40% of the total area, are quite dark in colour, and are heavily cratered. No other significant features are seen in these dark regions except for some vague furrows, probably remains of early giant impacts. The density of craters in these dark regions makes it clear that the crust here is roughly as old as that of Callisto. In contrast, the remainder of the crust is much less heavily cratered, and thus must have been remade long after the early era of bombardment. This younger surface displays terrain covered with a complex network of grooves or furrows (Figure 11.10). A major question has been to guess whether these grooves were produced by cryovolcanism (i.e. lava flows of liquid water rather than liquid silicate rock) or by tectonism (i.e. stretching and distorting of the surface by tidal or other stresses). Since almost no direct evidence of cryovolcanism (volcanic mounts or vents, flow fronts) are found, it seems clear now that most of the features on the grooved terrain were produced by repeated fracturing and stretching of the crust, probably as a result of the effects of large tides at some point in Ganymede's history.

The smallest Galilean moon, Europa, also presents a unique surface. From a distance, the moon's surface is extremely flat (relief of only hundreds of meters) and shows a global system of dark lines rather like a string wound around a ball (Figure 11.12). Only a few craters are visible, and it is clear that the present surface is very much younger than the moon as a whole. The models of interior structure indicate that although the moon contains only a small fraction of ice, that material has floated to the surface and overlies the much larger rocky and metallic interior. This is confirmed by strong evidence for the present of water ice in the spectrum of sunlight reflected from the surface. The surface of Europa thus is covered with a layer of H_2O approximately $1 - 200$ km deep. In recent years it has begun to appear likely that this layer may actually be mostly liquid: Europa may have a global ocean covered with an icy surface only a few km thick, as we will discuss below.

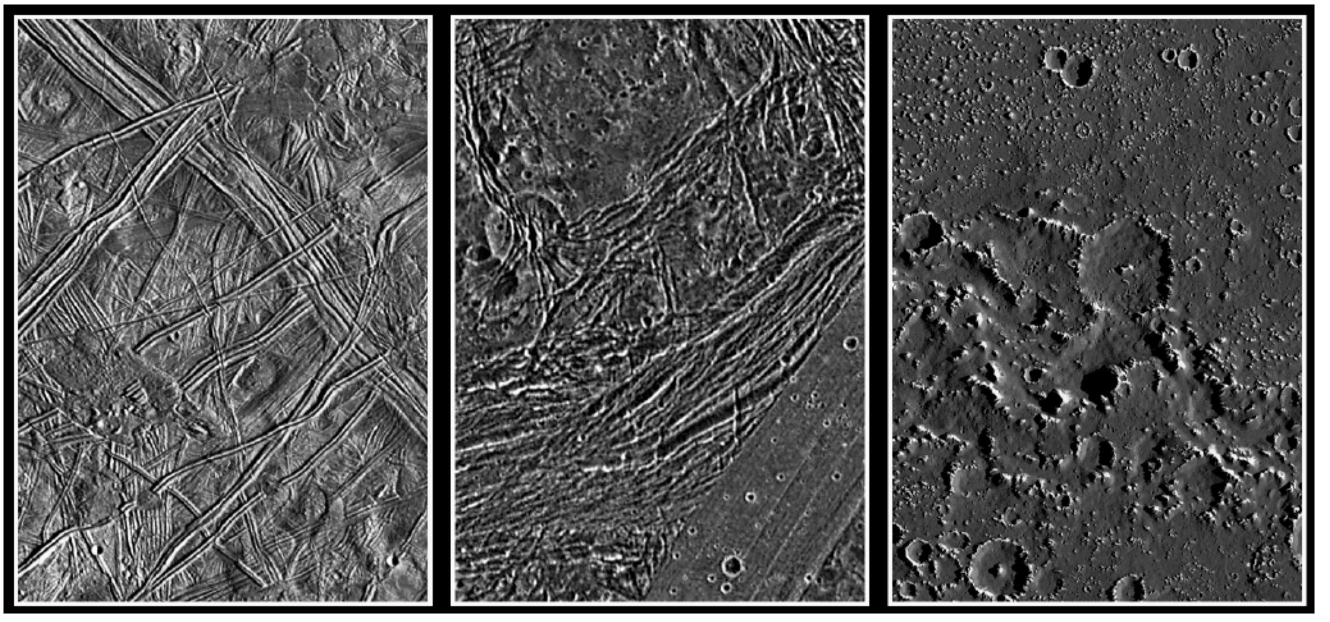

Figure 11.10: This figure shows small regions, about 100 km from top to bottom, on Europa (left), Ganymede (centre), and Callisto. Europa's surface is covered with a network of cracks and fractures, and appears to have been broken and re-cemented many times. No craters are seen. Ganymede shows some crates, but also much evidence of surface deformation and faulting. Callisto's surface shows nothing but craters interspersed with apparently smooth terrain. (Courtesy of NASA.)

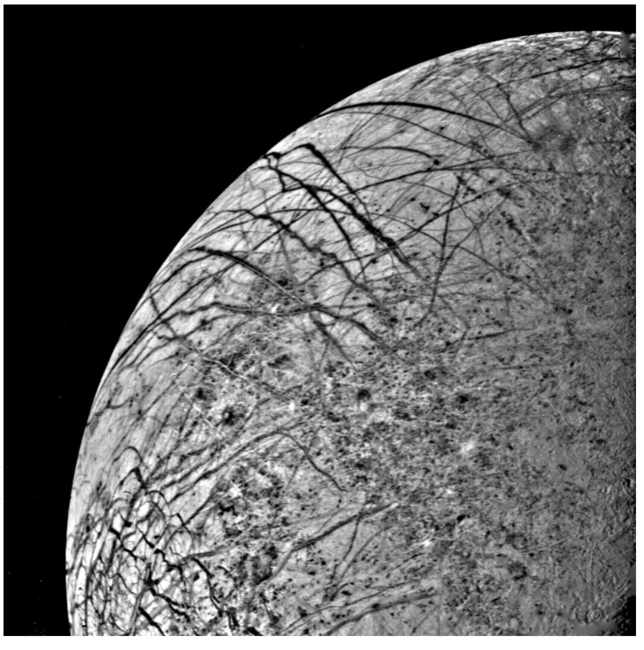

Figure 11.12: This image of about 1/4 of the the visible disk of Europa was obtained by the Voyager 2 probe in 1979. It clearly shows the dark lines criss-crossing the surface of the moon, and the absence of significant craters. (Courtesy of NASA.)

More detailed images from the Galileo orbiter reveal two dominant terrain types. Large regions on the moon are relatively smooth plains, on which the main visible features are long ridges of varying degrees of complexity. These ridges may be as much as a few hundred km long, and vary considerably in complexity. One common type is only about one km wide and about $1 - 200$ m high, and is split along its main axis by a steep central valley. Several of these long ridged plains (which are coloured by contaminants that are not yet identified) are visible in the left panel of Figure 11.10. These long ridges appear to have been created by some process of local crustal stretching and cracking.

A second major terrain type is known as "chaos". Such a region is seen to the left of the centre of the expanded view of Europa in Figure 11.13. These seem to be regions of the ridged plains that have been broken into pieces by local heating from below, perhaps set adrift in liquid water, and then re-frozen into a chaotic jumble of fragments.

It is widely suspected that Europa's icy crust may be liquid starting a few km below the solid surface, although this is not yet certain. One kind of evidence supporting this idea is the appearance of the chaos regions, which certainly appear to have existed at some point as icebergs in a liquid sea. Another line of evidence is the shallowness of even the largest craters on the surface. The recent impact crater Pwyll (which is

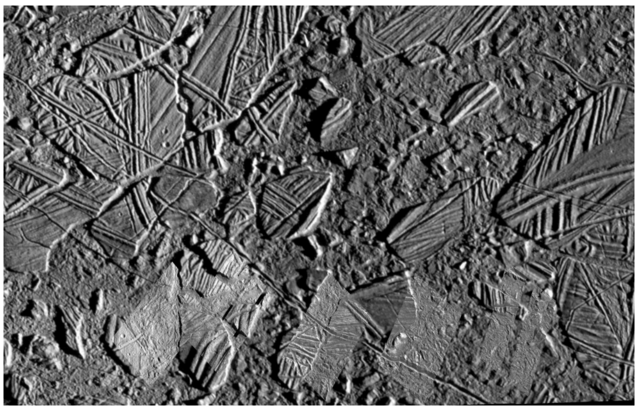

Figure 11.13: This image shows a small chaos region of Europa, about 35×50 km in extent. The surface is illuminated from the right (east). The crust here is clearly made up of fragments of plains terrain which have been broken apart, jumbled, and then re-cemented by introduction of (probably liquid) water in between the fragments. This region looks like what would happen on earth if a region of ice floes or icebergs in the sea froze over completely. (Courtesy of NASA.)

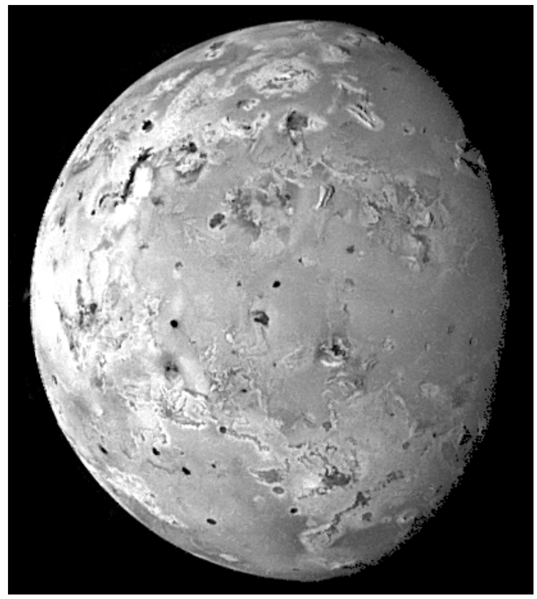

Figure 11.14: This black and white image of Io shows the patchy nature of the surface, but hardly does justice to the profusion of colours seen in the colour images of the moon. The image certainly gives the impression of a body on which lava flows have spread, and many of the dark spots appear to be lava sources. On the right side of this Galileo image, the boundary between light and dark is the terminator, where the solar illumination of the moon ends; near the terminator, several mountains cast sharp shadows. (Courtesy of NASA.)

probably no more than a few million years old) has a diameter of 26 km, but is only about 200 m deep. For comparison, similar sized craters on Ganymede (which also has an icy crust) are more than 2 km deep. One possible explanation for the remarkably shallow form is that the impact on Europa penetrated through a thin crust to liquid water, which then filled in most of the basin. Another possibility is that the surface has slumped and flowed like a terrestrial glacier can – but this would require that the ice temperature be close to the melting point, which would appear to require that the temperature below the crust is quite a lot warmer than the surface, and again strong suggests the possibility of a liquid ocean under the icy crust. NASA is currently considering a mission to the Jupiter system specifically to find out whether Europa does actually have an ice-covered ocean.

The innermost Galilean moon, Io, is a body which is completely unique in the solar system. The first images from the Voyager probes revealed a body resembling a pizza pie in colour, with irregular regions of white, yellow, red, brown and black on a surface pock-marked with spots and blotches. Closer inspection quickly revealed two remarkable facts: first, that the surface has no recognizable impact craters at all, and thus must be very young, and secondly that the moon has a large number of volcanic structures, several of which are actively ejecting lava and/or venting gases into impressive high plumes (Figures 11.14 and 11.15). It quickly became clear that Io is the most volcanically active body in the solar system.

Io has been further investigated using terrestrial telescopes and the Hubble Space Telescope, and also by the Galileo orbiter. These detailed investigations have confirmed the complete absence of detectable craters, a fact which implies that most of the surface of Io is renewed or replaced within about one million years by emplacement of new lavas. The rate at which the surface is covered by fresh lava is estimated to be of the order of 1 m per century on average! The detailed images have also revealed a large variety of volcanic landforms, including active lava lakes, lava flows and huge plumes, volcanic calderas, plateaus and plains. There are also impressive mountains (the highest known rises 16 km above the surrounding plains).

The colours revealed in the Voyager images strongly suggest that the lavas are rich in sulphur (many of the colours mimic those seen in the classic chemistry lab experiment of heating a test-tube containing elemental sulphur). Sulphur compounds are also detected in the spectra of reflected sunlight. These facts led to suggestions that the lava released from Io's interior might be largely sulphur, so that the entire surface

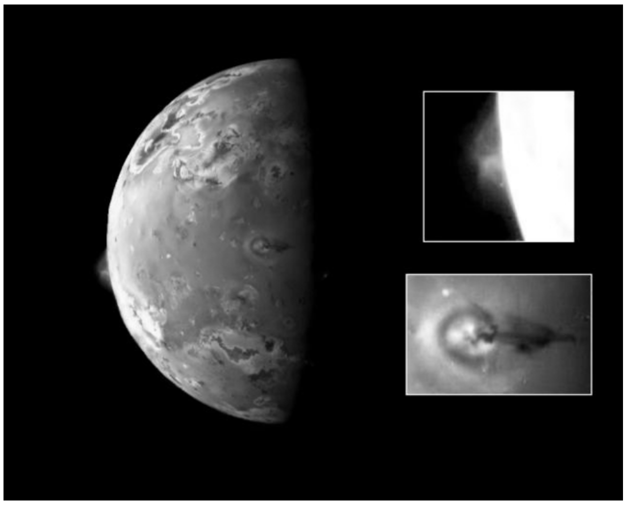

Figure 11.15: This image of Io is similar to the previous one, but reveals in two small close-ups volcano spewing gases to a height of well over 100 km. One volcano is silhouetted on the limb of the moon (left edge of the large image, and upper inset), and the other is seen from above, near the terminator (near the right edge of large image, about halfway down, and lower inset). (Courtesy of NASA.)

could be predominantly composed of sulphur. However, temperature measurements from the Galileo orbiter of the sources of lavas range up to 1700 K, far hotter than temperatures required to melt sulphur-rich lavas, and now it is thought that the lavas may well be magnesium- and iron-rich silicates like those in the Earth's mantle which are simply coloured by the addition of a small fraction of sulphur.

A body which is constantly pouring lava onto its surface can hardly be cool enough inside to have water trapped completely in the interior, so if there is any water in Io, some should be on the surface. The surface appears completely dry, however, and no spectral evidence of water is found. The moon has probably lost into space what little water it originally contained, as well as what it later accreted from impacts of comet nuclei.

Exercise: Why does the Earth's Moon not exhibit as much volcanic activity as Io does?

Origin of the moons of Jupiter

The observed properties of the moons of Jupiter raise a number of very interesting questions about how the system originated and developed. Why are all the inner moons in (almost) circular, coplanar orbits in the planet's equatorial plane? How did the three inner Galilean moons come to have orbital periods in the approximate ratios 1:2:4 – surely they were not formed by chance at just the right relative distances from Jupiter? And finally, why do these four bodies have such very different compositions and surface appearances, ranging from volcanic Io, a completely rocky body, to almost unaltered Callisto, which is about half ice and half rock?

Quite a lot about the moons of Jupiter can be understood by looking at three particular aspects of the moon system: how the moons formed in the first place, how they are heated internally, and how they have interacted through tidal effects with Jupiter and with each other since then.

The facts that the eight known inner moons have nearly circular, coplanar orbits close to the plane of Jupiter's (rotational) equator is strongly reminiscent of the situation of the planets around the sun. Like the moons, the planets orbit in nearly circular, coplanar orbits close to the equatorial plane of the sun. The reason for this regularity, in the case of the planets, is that they initially formed by condensation out of a disk of gas and dust that orbited around the sun as it accreted matter from its primordial cloud. Recall that this disk occurs because much of the material that is falling in towards the proto-sun is also falling *around* it (that is, the material has angular momentum). Infalling gas and dust clouds collide with the disk that is already present and join it. Material close to the sun in the disk is slowed down by friction with slower-moving gas and dust further out, and the inner material spirals gradually in to the central sun. As material condenses out of the disk to form first planetesimals and then planets, these are left orbiting the sun in the plane of the disk, and since the sun acquires most of its material from the disk rather than by "direct hits", the sun's equatorial plane ends up coinciding also with the disk.

The fact that the inner moons all orbit in the planet's equatorial plane strongly suggests that a similar process occurred on a smaller scale as Jupiter accumulated from interplanetary material. In fact, this is exactly what one would expect. In contrast, if the moons were formed separately, somewhere else, and later captured by Jupiter either through drag they encountered from the giant planet's initial, very extended atmosphere, or through interactions or even collisions with bodies already orbiting Jupiter, we would definitely *not* expect to find such a tidy, coplanar moon system.

The densities of the Galilean moons decrease systematically from Io (which has essentially no component of ice matter) to Ganymede and Callisto, both of which are about half ice, half rock and metal. This feature of the moon system is also reminiscent of the solar system as a whole, and confirms our suspicion that these moons formed in a proto-planetary disk. Recall the systematic changes in composition and density in the solar system as a whole, from terrestrial planets, whose composition is dominated by rock and metal, to giants and their moons in which ice is thought to be present in about equal amounts to rock and metals. These variations are explained by the different temperatures that reigned in different parts of the solar nebula, so that ice could condense and form part of the initial planetesimals in the outer solar system but not in the inner. Apparently a similar temperature variation existed in Jupiter's proto-planetary nebula, leading to two inner moons composed mainly of rock and metals, and two outer ones in which ices are an important constituent. Thus, we believe that the overall orderliness of the inner, regular moons, is due to their formation in a proto-planetary disk around the forming planet.

In contrast, the two outer systems of irregular moons are probably the result of capture of one or more stray planetesimals by Jupiter. This would be possible either due to the drag exerted on the initial bodies by an early extended atmosphere, or due to collisions of the initial bodies with other orbiting debris. Quite possibly each of the two presently known irregular groups of moons originated with a single body that has been fragmented and perhaps re-assembled as a result of collisions with other orbiting material, or with comets passing through the Jupiter system.

Thermal evolution of the Galilean moons

Let us now return to the Galilean moons, and try to understand their histories. As these moons formed, each released a rather large amount of gravitational energy – infall motion that was converted to heat as chunks of material fell onto the proto-moons. Enough energy was potentially available from this source to raise the average interior temperatures (from initial values of 100 or 200 K) of Ganymede and Callisto by several hundred K, of Europa by roughly 1500 K, and of Io by nearly 3000 K. Of course, some of this heat was radiated back into space as the moons formed. How large a temperature rise each moon actually achieved depended on how quickly each formed (quick formation would have given little time for heat loss by radiation and so would have led to high internal temperatures), and also on the details of how the material fell onto the moons, neither of which is known.

Since the energy released while the proto-moon is small is also rather small, the heating effect is not significant until the proto-moon has reached a radius of a few hundred km. Beyond that size the heating rapidly becomes more important (unless accretion is very slow, so that most of the heat released can be radiated away as it is released by debris impacts). Thus we expect that the heat deposited in the moon by the accretion process will be distributed quite unevenly. The temperature rise in the outer layers can easily be as large as the estimates given above, but in the deep interior of the forming moon the temperature would have initially been not much higher than the temperature of the nebula in which the moon formed, perhaps in the range of 100 to 200 K.

The heating of the outer layers to a temperature only 200 or 300 K above the nebular temperature would have been sufficient to melt the ice component of the moon outside of a radius of somewhere between 1000 and 2000 km. We would expect this to have occurred in the large icy moons Europa, Ganymede and Callisto, as well as in Saturn's Titan, and possibly in Neptune's moon Triton. In the outer layers of these large moons, then, the rocky component would have settled to the bottom of the melted layer, while water ice and the volatile materials absorbed in it would have risen to form a thick low-density layer which was initially liquid but which would have frozen – at least near the surface – rather quickly. Each of the large moons icy would have found itself with a rather strange density profile: a core of undifferentiated rock and ice with a density of about 2000 kg m^{-3}, then a silicate layer with a density closer to 3000 kg m^{-3}, and finally an ice and water layer of density around 1000 kg m^{-3}. Such a situation is quite unstable, but as long as the moon's core was quite rigid, it would not have changed.

However, the silicates of the core contain naturally radioactive substances just like the materials of the terrestrial planets (though somewhat diluted by the ices). Heat released from this source would have gradually heated the moon's mixed rock and ice core until the core material was soft enough for the denser layer of silicates to sink into the core, and for the water ice in the core to escape into the mantle. This event is known as **core overturn**, and would probably have occurred roughly 1 Gyr after formation of the moon. As a result, the moon would have developed a rocky core, surrounded by a mantle of water and ice, possibly mixed with other volatiles such as ammonia, methane, nitrogen, and carbon monoxide.

How the core of the moon would evolve after this point depends on another unknown of the story – the initial chemical composition of the rocky component of the moon. At one extreme, we might imagine that the rocky material in the moons of the outer planets resembles CI meteorites, with *both* the silicates and the iron highly oxidized and combined chemically in pyroxene and olivine minerals. In this case, the core would continue to heat up, losing heat at the same time by conduction into the surrounding, cold water- and ice-rich mantle. Eventually the core would begin to slowly turn over by solid-state convection like that which occurs in the earth's mantle (see Chapter 8), but this would not greatly alter affairs. The temperature of the core would now probably be somewhat above 1000 K.

At the other extreme, the moon could have formed with a rocky component similar to the composition to the EH (enstatite) meteorites, in which iron is almost entirely in metallic form. (The material that formed the earth may have had roughly this composition.) In this case, the gradual heating of the rocky core would have eventually raised the temperature (1 to 2 Gyr after core overturn) to a high enough value to melt the iron and differentiate the core. In this case the moon would end up with a three-layer structure of a liquid iron inner core, a solid silicate outer core, and and ice and water mantle. Again the core would presently have a temperature above 1000 K, and slow convection would occur in the rocky mantle.(Because the gravity measurements of the Galilean moons by the Galileo probe have not been quite sensitive enough to reveal possible iron cores, we do not yet know whether these moons have iron cores or not.)

In the icy mantle, the situation also depends on the unknown chemical composition. Here, the key unknown is the amount of volatiles other than water the layer contains. This is quite important because substances such as NH_3 can act as powerful anti-freeze agents. The melting temperature of water depends on the pressure, but never drops below about 250 K (-25 C). However, with a few percent of ammonia dissolved in the water, the melting point is depressed to 175 K (about -100 C). This drastically alters the situation in the outer ice-rich layer.

When the moon forms an icy mantle as a result of core overturn, this mantle is initially probably liquid throughout. However, the outer surface of the moon quickly cools to the temperature set by incoming sunlight, around 120 K (-153 C) for the moons of Jupiter. At this temperature the surface freezes solid, and a surface layer of solid ice develops like the ice cover on a lake in winter. What happens next is still uncertain, both because we do not know what substances (besides water) are present in the mantle, and because the behaviour (particularly the conditions for solid-state convection to occur) of ice at low temperature is still not fully understood. But we can roughly bracket the possibilities by looking at extremes.

One possible evolution occurs if the ice mantle contains no important substance that can act as an antifreeze, and if solid-state convection (like that in the Earth's mantle) can occur in the ice where it is not too cold. In this case, the surface temperature of the solid surface ice layer overlying the deep sea is held at the equilibrium temperature set by sunlight, but the bottom layer of the ice layer is at the melting point of water, which is between about 260 and 280 K at various depths. The temperature *difference* across the ice layer is thus held roughly constant at about 150 K, and the inner part of the ice sheet is at a temperature not far below the freezing point. This may make it possible for the lower part of the solid ice lid to convect slowly. This is a very efficient mechanism for carrying heat outwards towards the surface, much more effective than simple conduction of heat through a layer tens of km thick. In this case, the heat transported to the surface per second is easily *larger* than the heat released by radioactivity in the silicate core, so more heat flows out of the mantle from the top than flows in from the bottom. The mantle cools rapidly, and within one or two Gyr the mantle is completely solid ice. From that time to the present, this mantle has been in a state of solid-state convection, simply carrying to the surface the heat released by radioactivity in the core. As the radioactivity level decays, the core and mantle both gradually become cooler.

As the other extreme, we consider the possibility that there is an important antifreeze (such as NH_3) in the material of the mantle. Ammonia is capable of lowering the freezing temperature of the sea by as much as 100 K, to about 170 K. In this case, the top of the surface ice sheet is held at about 120 K, while the bottom is at the melting point of the liquid, only about 50 K warmer. In this case, it seems very likely that convection in the lower part of the surface ice sheet will be prevented by the very low temperature, and with only a small temperature drop across the surface ice sheet, heat conduction out to the surface can easily fall, as the surface ice sheet thickens, to a rate that balances the heat production in the silicate core. In this case, once the surface ice sheet (the lithosphere) becomes thick enough (a few tens of km) for heat loss from the surface to balance production in the core, the low-density mantle almost ceases to cool further. In this situation, a deep liquid sea could still exist today under the solid surface even in a moon such as Callisto that has no significant internal heat source except for radioactive heat release.

If the moon has no important anti-freeze substance in the mantle, but solid-state convection in the ice lid does not occur (there is still dispute on this point), then the lid can still gradually grow to such a thickness that the heat carried out through the ice sheet by conduction falls to a value that balances production in the core. Again, a deep liquid ocean covered by a solid lithosphere of ice would persist up to the present. Thus, it is quite unclear at present whether the large Galilean moons Ganymede and Callisto have liquid seas under their icy surfaces, but it is not unreasonable to speculate that they do.

From this general picture, we would predict that all the Galilean moons, even rocky Io, would have differentiated and developed solid lithospheres billions of years ago. The three ice-rich moons might still have liquid seas deep below the surface, but their surfaces should all be primarily composed of water ice, and heavily cratered.

For Callisto, we indeed find the cratered surface, but recall that the gravity measurements show that this moons is *not* fully differentiated into a silicate core surrounded by an icy mantle, although "some" differentiation seems to have taken place. Furthermore, the surface is not pure ice, but has a lot of dark, probably rocky, material in it. We are thus left with a rather serious puzzle as to how Callisto could have avoided complete differentiation, and how the surface could remain a mixture of ice and rock.

Ganymede has a nearly pure ice surface, and gravity measurements show that it has indeed fully differentiated into a core-mantle structure, as expected. However, recall that the surface shows both regions of heavy cratering, and other that have been heavily modified by tectonic events after the end of the main period of bombardment. Again we have a puzzle.

Gravity measurements of Europa confirm that it is differentiated, and its surface appears to be pure ice. However, it has virtually no craters on the surface, so some effect has reworked the entire surface within the past few millions of years. Io has a surface which does not have a single impact crater, and is covered with volcanic structures and other evidence of tectonic activity. Clearly there is some important part of the story we have so far neglected. This is probably the effect of intermittent heating due to orbital resonances.

Orbital resonances and tidal effects

It appears that both the orbital resonances that are found for the three inner Galilean moons, and the fact that all of their surfaces show evidence of melting long after the formation period, may be due to the effects of tides. Let us recall what kinds of effects might me expected. First, if the moons initially formed with rapid rotation (periods of a few hours, perhaps) around their rotation axes, Jupiter would produce tidal bulges on the near and far sides of each moon. These bulges would be pulled ahead of the Jupiter-moon line by the rapid rotation, and so Jupiter would exert a drag on the rotation of each moon (just as the Moon exerts a drag on the rotation of the Earth). As a result, each of the Galilean satellites would slow down to synchronous rotation within a few million years. In fact, they are all observed to rotate synchronously. While this slowing of the rotation was occurring, each of the moons would be dissipating energy in its interior – there would be strong *tidal heating*. In contrast, the outer small moons are all too far from Jupiter for tides to be effective, and they are not rotating synchronously.

Another tidal effect is that is important is due to the pair of bulges that each moon would raise on Jupiter. Jupiter's rotation about its axis occurs with a shorter period (presently 0.41 d) than the orbital periods of any of the Galilean moons, so the bulges would be carried ahead of the Jupiter-moon line. The extra attraction for the near bulge by the moon causing it would have the effect of slowing (slightly) the rotation of Jupiter, while the bulge itself gives the moon a bit of extra pull, gradually increasing the orbit size (and angular momentum) of the moon. This effect causes the orbits of each of the moons to gradually increase in size, but since the effect falls off rapidly with distance from Jupiter, orbit growth occurs more rapidly for Io than for Europa, and more rapidly for Europa than for Ganymede. This effect has been occuring throughout the history of the Jupiter system, so the moons' orbital sizes are larger than they originally were. Again this is reminiscent of the situation of the Earth and Moon.

The gradual expansion of satellite orbits, with the innermost ones growing most rapidly, has another very interesting effect. As the orbit of Io expanded from it initial rather smaller size, eventually Io came into an orbital resonance with Europa, perhaps a 2:1 resonance like the present one. It appears that the two moons are likely to become trapped in this resonance, so that as Io is pushed farther and farther from Jupiter, it pushes Europa out in front of it, always keeping the 2:1 relationship between the two orbital periods. Eventually, both orbits expand enough that the orbit of Europa reaches a 2:1 resonance with Ganymede. Again trapping occurs. Thus, we do not imagine that the Galilean moons originally formed with orbital period ratios of approximately 4:2:1, but rather evolved into this state. (Eventually all three orbits will expand enough to reach resonance with Callisto too, but this is far in the future.)

As the orbital periods of the two moons approach such a resonance, the regularly occuring pull that each feels from the other causes both orbits to develop a significant eccentricity. That is, the effect of the regularly occurring mutual attraction is to cause each of the moons to vary its distance from Jupiter. Since the amount of tidal stretching of each moon depends on how far it is from Jupiter, the tidal distortion of the shape of each moon varies – it stretches and relaxes, stretches and relaxes – as it goes around in its orbit.

This periodic change in the shape of the moon is a dissipative process – it deposits energy into the interior of the moon. How much? Roughly the correct amount to explain the heating that must be occurring inside Io to produce the ongoing volcanic activity that we observe. And possibly enough energy is dissipated inside Europa to lead to an ocean below the icy surface. This effect may even have been sufficiently important for Ganymede at the time that it came into resonance with Europa to explain the regions of deformed crust that are not seen on Callisto.

We do not yet have answers for all the interesting questions about the Galilean moons (for example, why are Europa and Ganymede so different in surface appearance?), and some of the ideas discussed above still have important uncertainties. However, it appears that quite a lot of the essential physics of this fascinating moon system is gradually becoming clearer.

The moons of Saturn

In contrast to the moon system of Jupiter with its four large (but very dissimilar) moons accompanied by a dozen or more very small bodies, the moon system of Saturn is dominated by a single large moon, Titan. With a diameter of 5150 km, about 1.5 times larger across than the Earth's Moon, Titan is the second largest moon in the solar system, after Ganymede. Orbiting the ringed planet together with Titan are seven smaller but still significant moons having a variety of sizes: two (Iapetus and Rhea) with diameters near 1500 km, two (Dione and Tethys) with diameters of about 1100 km, two with diameters of around 450 km (Enceladus and Mimas, which are spherical), and one about 300 km (Hyperion, which is potato-shaped). Moons of this intermediate size are not found in the Jupiter system, and it is very interesting to examine them to discover how they differ from moons of the size of the Galilean satellites. An intriguing feature of this group of intermediate-size moons is that the five that orbit inside the orbit of Titan are found in decreasing order of size as one goes inward.

Finally, there are at least ten smaller moons, mostly not not spherical, with characteristic sizes of 250 km or less. The smallest moons known in the Saturn system are only about 20 km across. Unlike the tiny inner moons of Jupiter, which appear to be rocky (as one would expect for bodies formed in the inner part of a warm proto-planetary disk), the smallest moons of Saturn appear to be rich in ices.

The regular structure of the moon system around Saturn strongly suggests that the moons formed in an equatorial disk of material that was accreting onto the main planet, as we have deduced for Jupiter. The fact that all the moons of Saturn have relatively low density, and thus contain a major component of ice, indicates that the accretion disk around Saturn was substantially colder than that of Jupiter, so that more volatile substances could freeze out and be swept up by the forming moons, even close to the planet.

The inner satellites of Saturn, especially from Dione inwards, have been strongly affected by collisions with passing comets. Comets that happen to pass near Saturn are of course strongly attracted by its gravity, and for this reason make a closest approach to the planet that is considerably closer than it would be if Saturn were not attracting them. This effect increases the likelihood of a collision between such comets and the inner moons of Saturn, as well as the typical speed with which the comet impacts the moon. As a result, it is probable that all of the moons from Dione inwards have suffered at least one impact carrying enough kinetic energy to leave a crater with a size comparable to the impacted moon. In many cases these impacts would have been powerful enough to disrupt the satellite, which could later have re-accreted into one or more bodies.

The moons out as far as Hyperion (which is slightly outside the orbit of Titan) appear to have undergone increases in the sizes of their orbits due to tidal effects, just as happened in the Jupiter system. As a result, several orbital resonances have been established. The orbital period of Tethys is twice as long as that of Mimas, as is true of Dione and Enceladus. These 2:1 resonances have been plausibly accounted for as a result of more rapid tidal expansion of the inner orbit than the outer, until the two moons "lock into" a resonant situation. Farther out, Hyperion makes three trips around the planet four every four made by Titan, but the tidal evolution that led to this situation is quite unclear. There are also at least three situations of shared orbits: Janus and Epimetheus, Telesto, Tethys, and Calypso, and Helene with Dione; some of these may be the result of the disruptive collisions mentioned above.

These tidal effects have also strongly affected the rotation of Saturn's moons. All the moons out to Titan are locked by tidal forces into synchronous rotation; each keeps one hemisphere constantly facing Saturn.

Titan (Figure 11.16) is the only moon in the Saturn system which is comparable in size to the Galilean moons, Neptune's Triton, and our own Moon. Because of its size and composition, it is very likely that Titan has undergone an evolution similar to that which has apparently occurred inside Ganymede. Accretion heating probably melted the outer layers of the moon as it formed in the disk of material orbiting the equator of the forming planet, resulting in the formation of a layered moon with an undifferentiated core and a mantle with a dense silicate layer beneath a layer of

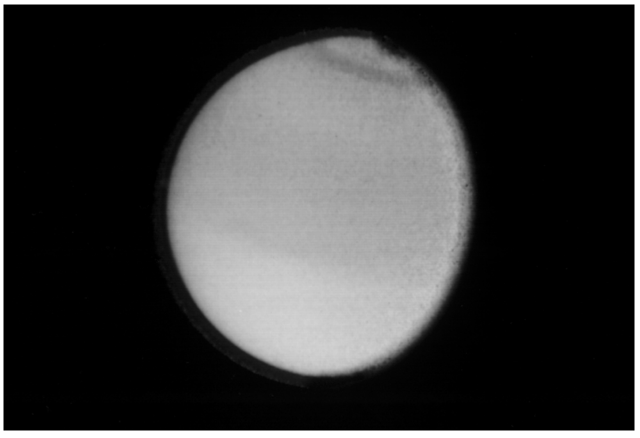

Figure 11.16: Saturn's only really large moon, Titan, is a body composed about half of rock and half of ices. It is probably similar to Jupiter's Ganymede in structure. It has a nitrogen-rich atmosphere; a haze of aerosol particles in the atmosphere makes observation of the surface very difficult. (Courtesy of NASA.)

water and ice. As the core warmed up from radioactive energy release, ice in the core softened, melted, and rose towards the surface, while the dense silicate layer above sank into the core. By about one Gyr after Titan first formed, core overturn was complete. In the mantle, now composed primarily of water and perhaps other volatiles, we again have a competition between the heat that flows into the mantle from the core below, and cooling through the outer icy layers to the frigid surface. As with Ganymede and Callisto, we are not sure how efficient heat loss out through the lithosphere is, so we do not know whether Titan has an ocean deep below its icy surface, or whether the mantle is frozen solid throughout. The Huygens–Cassini space probe mission, which should arrive at Saturn in 2004, will probably help greatly to understand the structure of Titan.

However, Titan is also entirely unique in that it is the only moon in the solar system with a massive atmosphere. Unfortunately, this atmosphere is filled with a haze of small particles (Figure 11.16), and so we cannot see the surface in visible light. Our information on this singular gas envelope comes from observations by both Voyager probes (Voyager 1 was sent quite close to Titan), and from a range of observations from the earth. By observing the way in which the Voyager radio signal changed as the craft passed behind Titan, together with infrared brightness measurements, it was found that the mean molecular weight of the gas making up Titan's atmosphere is close to 28, about the same as that of molecular nitrogen, N_2. The surface pressure is about 150 kPa (1.5 bar), 1.5 times the surface pressure on Earth. Since gravity on Titan is only about 13% of that on Earth, there is about 10 times as much mass of gas over each square m on Titan as on Earth!

The surface temperature is found from infrared observations of brightness to be about 94 K (about −180 C). The temperature declines slowly with height to a low of about 70 K at 30 km. Above this level, the thermosphere begins (Titan has no stratosphere like that of Earth), and the temperature rises to a balmy 170 K (−100 C) at around 200 km above the surface. The atmosphere appears to be convective near the surface.

The chemical composition of the atmosphere is dominated by N_2 (probably about 94% by mass), with a few percent of CH_4 as the principal minor constituent. There are also very small amounts of CO, CO_2, C_2H_6 (ethane), and a number of other compounds of H, C and N. The predominantly nitrogen composition is surprisingly similar to the atmosphere of Earth.

The presence of CH_4 in Titan's atmosphere presents an interesting puzzle. It is found that this gas is rapidly broken up by energetic photons from the Sun into C (which combines with other molecules in the atmosphere to form substances that settle on the surface) and H (which escapes from the moon because of its feeble gravity). All the methane in the present atmosphere would be destroyed in this way in about 30 million yr. This strongly suggests that there is a continuing source that replenishes this molecule as it is destroyed. It is thought that this could be in the form of volcanic activity releasing CH_4 from the interior of the moon – if the interior is hot enough – or in the form of methane lakes on the surface.

How did Titan come to have an atmosphere when none of the Galilean moons acquired one? We can be sure that the moon did not acquire its atmosphere by direct gravitational capture from the proto-planetary nebula in which it formed because in that nebula the abundance of neon was about the same as that of nitrogen, but the current atmosphere of the moon has less than 0.1% of Ne. It also appears that the bulk of the atmosphere was not delivered as a result of cometary impacts, because the ratio of normal hydrogen to deuterium (hydrogen with an extra neutron in the nucleus) is rather different in comets than in Titan's atmosphere. So it appears that the difference between Titan and the Galilean satellites is that Titan acquired a good stock of the gases now found in the atmosphere when it formed, while the Galilean moons did not.

If the gases of the atmosphere were acquired at the time of formation, they were presumably trapped inside ice grains that formed the planetesimals that became Titan, as clathrates. In particular, ammonia (NH_3) is easy to trap in this way, and a significant supply of (more volatile) CH_4 might have been acquired in the same way. It appears that the key difference between the nebulas of Jupiter and of Saturn was probably that temperatures in Saturn's proto-planetary nebula were substantially lower than in Jupiter's, leading to a much greater trapping of volatile molecules in the icy materials that went into forming Titan. The Galilean moons lack atmospheres because the nebula in which they formed was too warm for significant inclusion of suitable volatiles in the ices that formed them.

But was the nitrogen that makes up Titan's atmosphere originally trapped as NH_3, or as N_2? Molecular nitrogen, N_2, was probably the dominant form of N in the *solar* nebula, but could have been converted to NH_3 in the Saturn proto-planetary nebula. If we look at the ratio of current nitrogen in the atmosphere to current argon, which would have been trapped in ices in about the same proportion as N_2, we cannot come to any particular strong conclusion, since the atmosphere of Titan may contain up to about 6% Ar. However, looking at the ratio of ^{15}N to ^{14}N, which is a good indicator of how much N has *escaped* from the atmosphere (remember, the heavier atoms escape less easily than the light ones), we conclude that the early nitrogen atmosphere of Titan was probably about 30 times as massive as the present one! Then the deduced ratio of total nitrogen to observed argon (which is too heavy to escape from the atmosphere) provides a strong argument that the nitrogen was *not* accreted by Titan as N_2. It must have been originally trapped as NH_3, which is quickly converted by solar ultraviolet light to N_2 when it reaches the atmosphere from the interior.

We are still rather unclear about how the CH_4 in the atmosphere was acquired. Planetary scientists eagerly await the arrival of the Huygens–Cassini space probe in early 2004; this probe should help to clear away at least some of the present uncertainties surrounding this particularly moon.

We next turn to the smaller moons. Six of the intermediate moons have known masses (Iapetus, Rhea, Dione, Tethys, Enceladus, and Mimas). All have densities between 1440 and 1160 kg m^{-3}. These bodies are mainly composed of ice, though each must have some rock. No clear pattern of density decreasing with distance, like that observed for the four large moons of Jupiter, is seen here. All six moons show clear evidence, in the way they reflect infrared light between 1 and 2 μm, of water ice on their surfaces, as one might expect from their densities.

All of these moons (except for a part of Enceladus, discussed below) are heavily cratered, a fact evident in the Voyager image of Dione (Figure 11.17). The craters generally have more vertical relief (their walls are high above the central floors) than is found for Ganymede and Callisto. This is probably due to a more rigid crust

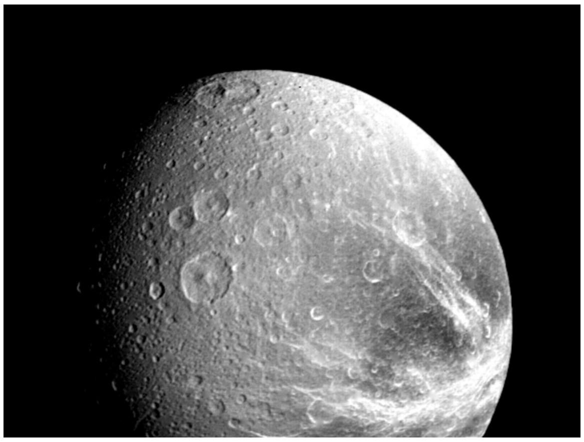

Figure 11.17: Saturn's intermediate-sized moon Dione, with a radius of 560 km, is large enough to be quite spherical. The surface is heavily cratered, with more vertical relief than is found on Jupiter's moon Callisto, perhaps because of the colder and more rigid crust of Dione. The strong variations in reflectivity are not yet understood. (Courtesy of NASA.)

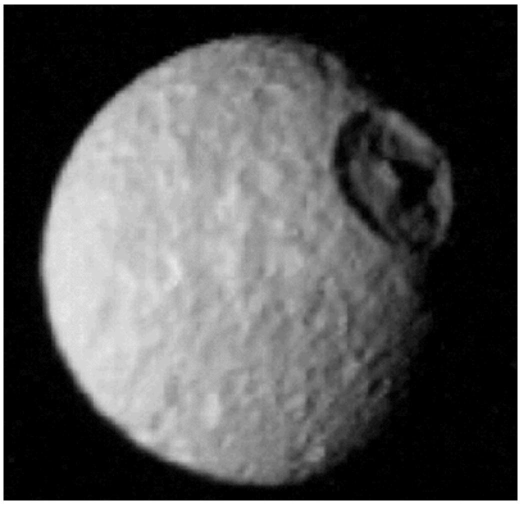

Figure 11.18: Mimas is the inner-most of the intermediate size moons of Saturn, and was not imaged from nearby by either Voyager. However, one huge crater, 130 km in diameter, is easily visible even in this low-resolution image. (Courtesy of NASA.)

on these small and very cold bodies. Impact craters that are so large that it is surprising that the moon survived are found; an impressive example is that of Mimas (Figure 11.18). Several of the intermediate moons show long chasms or cracks that might have been produced as a result of impacts.

Another remarkable feature of the surfaces of several of the moons is regions of rather low reflectivity. Iapetus is extremely dark (it reflects only 5% of the light striking it) on the hemisphere which faces forward in the moon's orbit around Saturn (the "leading" hemisphere), but is much brighter (an albedo of 50%) on the "trailing" hemisphere. It is thought that Iapetus may have swept up much dark debris broken loose by meteorite impacts from Saturn's very dark outermost moon, Phoebe, which is probably a captured carbonaceous asteroid. In contrast, both the middle moons Rhea and Dione (Figure 11.17) have streaky dark terrain at the centres of their trailing hemispheres, but are elsewhere quite reflective. These dark regions may be the oldest surviving surface on these two bodies; elsewhere the surface has been reworked by the impact of small meteoroids since the moons formed.

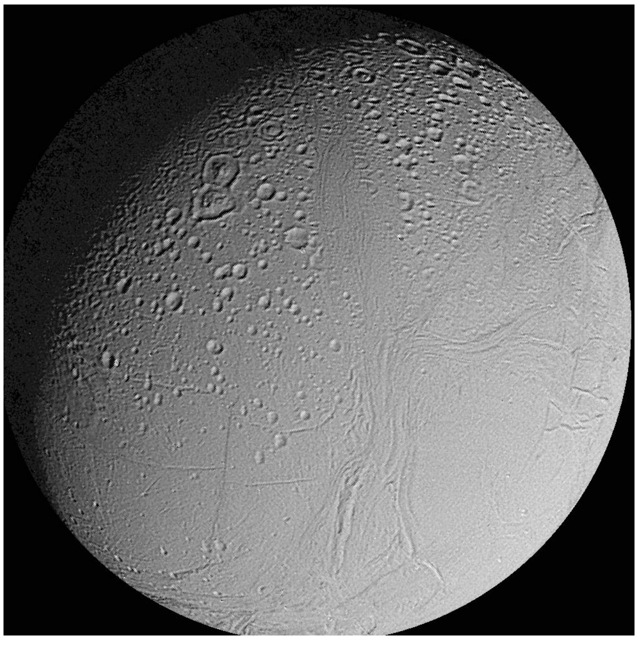

Figure 11.19: The small Saturnian moon Enceladus shows a remarkable variety of surface features, ranging from impact craters to the wide band of grooves that traverse the moon from upper left to lower right. This grooved terrain is reminiscent of that found on Ganymede, and clearly shows that this small moon has undergone important tectonic activity. (Courtesy of NASA.)

Finally, the small, highly reflective moon Enceladus (the second innermost intermediate moon, after Mimas) shows a remarkable variety of terrain, ranging from one hemisphere of heavily cratered terrain to regions which have been resurfaced so recently that they are completely free of craters, but instead are criss-crossed by a network of banded ridges (Figure 11.19). The resurfaced terrain is reminiscent of that on Ganymede. This crater-free hemisphere is almost 100% reflective, like fresh snow or pure, clean ice. The resurfaced hemisphere is estimated, from the absence of large craters, to be less than 1 Gyr old, so this area has apparently been resurfaced in the most recent 20% of Enceladus' history.

Understanding the energy source that provided the heat to melt water and resurface a part of Enceladus has proven to be very difficult. The moon is far too small to have heated significantly during accretion, and its low density shows clearly that it does not have a very large rock component, so that there is certainly not an important radioactive heating source inside (all the major radioactive substances are in the rock component). The only obvious source of heating is tidal friction, the effect that keeps Io's volcanos active. However, the present orbit of Enceladus has rather low eccentricity, and the estimated tidal heat input rate at present is about 300 times too small to provide the energy needed to resurface the moon. Furthermore, the nearby moon Mimas has a *more* eccentric orbit than Enceladus, but it shows few signs of recent surface melting. The energy input that made possible the reworking of Enceladus' surface was probably the result of some tidal resonance with another moon in the past, but the sequence of events that led to the heating is very unclear at present.

The bulls-eye moon system of Uranus

Before the Voyager 2 visit to the Uranus system in 1986, little was known about the moons beyond their orbital properties. Five moons were known, spanning the sizes of the intermediate-size moons of Saturn. All orbit close to the equatorial plane of the planet, in orbits located between about 5 and 23 Uranus radii from the planet's centre. All these orbits are direct, and the rotation of each of the moons is synchronized with the orbital period. These moons get steadily larger as one goes out. First comes tiny Miranda, with a radius of about 240 km, then Ariel and Umbriel, both with radii of about 580 km, and finally Titania and Oberon, with radii of about 770 km. These moons are all well outside the system of dark rings discovered from Earth.

One remarkable feature of the Uranus system is that the equatorial plane of the planet, which is also the orbital plane for the regular moons, is inclined to the ecliptic plane by 97°. Thus the "north" pole of Uranus (if you were above this pole you would see the planet rotating counter-clockwise) is actually slightly below the ecliptic plane. This is different from the other planets, all of which have the north poles of their rotation axes pointing roughly perpendicular to the ecliptic plane on the same side as the north pole of the Sun. Because the rotation axis of Uranus remains fixed in direction as the planet orbits the Sun, a unique seasonal pattern occurs on Uranus, and on its moons. As the system circles the Sun in its 84-year period of revolution, first one of the poles points towards the Sun for some years. The line to the Sun gradually shifts as the planet continues in its orbit, and 21 years later the Sun is over the equator (this will be the situation in 2006). Another 21 years brings the opposite pole to point nearly to the Sun, and so on. This arrangement has the consequence that one hemisphere is constantly illuminated for some years while the other it permanently dark. As the line to the Sun shifts towards the equator, the days begin to alternate like those on other planets, with the rotation period of the planet determining the length of the day for Uranus, and the orbital periods determining the length of days on the moons. As the planet-Sun line shifts towards the opposite pole, the situation reverts to permanent illumination, now of the opposite hemisphere than before.

Exercise: Sketch the Uranus system relative to the ecliptic plane and convince yourself of the correctness of the description above of the seasons. When was the last year in which one of the poles was pointing towards the Sun?

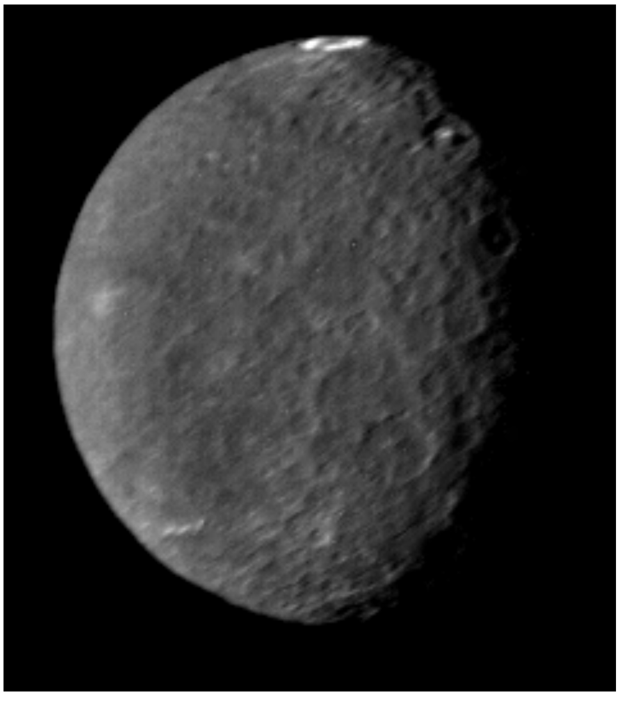

Figure 11.20: A distant view of Uranus' heavily cratered moon Umbriel. Little can be seen on this moon except impact craters; some of the larger ones have central mountain peaks. (Courtesy of NASA.)

Because of the inclination of the orbit of Uranus to the ecliptic, when Voyager 2 approached the system in 1986 the moon system resembled the bull's eye of a target. The space probe was able to obtain a close-up view of only one of the moons, Miranda, and although the moons all rotated during the time Voyager 2 passed through the Uranus system, only one hemisphere was ever illuminated. At least half of each of the moons has thus still not been photographed.

The Voyager visit made possible the accurate determination of the sizes and masses of all the previously known moons, and led to the discovery of another 10 much smaller moons, all orbiting within the orbit of Miranda in the planet's equatorial plane in nearly circular orbits. These bodies (which are mostly too small to reveal any surface details even to Voyager 2) have radii of between about 10 and 80 km.

Using the newly determined dimensions and masses of the larger moons, their densities are found to range from 1200 to 1710 kg m^{-3}. They are all somewhat denser than moons of similar size in the Saturn system; thus although they are – like the moons of Saturn – composed of a mix of rock and ice, there seems to be somewhat more rock in moons of Uranus than in those of Saturn.

All the moons show some cratering. Two, Umbriel and Oberon (the third and fifth of the intermediate moons, counting outwards) have heavily cratered surfaces, rather like those of Dione or Mimas (see Figure 11.20). The cratering is about as dense as that of the lunar highlands, and probably was produced during the period of intense bombardment that ended about 4 Gyr ago. These moons have apparently been largely inactive since that time.

Ariel and Titania (the second and fourth intermediate moons from Uranus) also have much cratered terrain, but there are few really large craters like those found on Umbriel and Oberon (see Figures 11.21 and 11.22). Since all four moons must have suffered similar bombardment by large planetesimals during the first half-billion years of their existence, the absence of large craters indicates that the surfaces of these moons must have melted or been covered with "lava" (liquid or slushy water, perhaps mixed with other volatiles such as ammonia) early on. The current crop of craters must date from after this time, and from after the end of the period of heaviest bombardment. In addition to the craters, each moon is deeply scarred by long trenches (grabens) where the crust has cracked open and been partly filled in by liquid or slushy solid from below. The origin of these long cracks is uncertain, but they appear to indicate that the interior of the moon expanded while the lithosphere did not. The result was long fractures in the crust. The most obvious way in

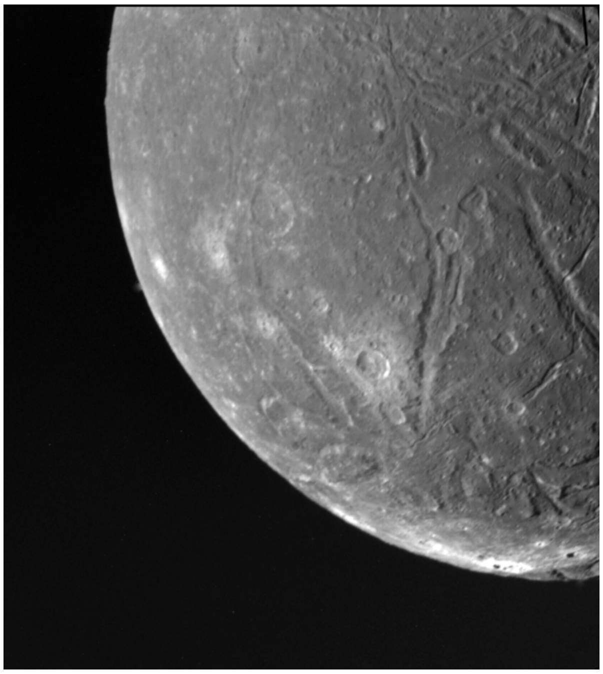

Figure 11.21: This view of Uranus' moon Ariel from Voyager 2 shows clearly that Ariel is cratered but not very heavily, and also that the moon has an extensive system of surface faults that appear to have been produced by expansion of the surface, probably caused by internal heating. (Courtesy of NASA.)

which this could happen would be if the interior were heated and therefore expanded while the outer layers remained cold.

Finally we come to the smallest and innermost of the major moons, Miranda. It was widely expected that this small body would be heavily cratered but otherwise quite uninteresting. In fact, the images sent back by Voyager 2 revealed what is easily the most puzzling surface of any of the moons of Uranus. Part of Miranda is indeed covered with rolling plains with some cratering, but other parts have largely uncratered oval or trapezoidal regions covered with wide grooves and ridges that seem to follow the outline of the region. A small part of Miranda which shows both kinds of terrain in seen in Figure 11.23. This groove-and-ridge terrain is quite mysterious; one possibility is that this moon has been completely fragmented by impact and then reassembled by gravity, and that the groove-and-ridge terrain is somehow the result of this process. Again we have strong evidence in the relatively young surface, as well as in the strange oval regions, of important resurfacing activity in the moon after the end of the heavy bombardment period.

We have seen in the Jupiter and Saturn systems that

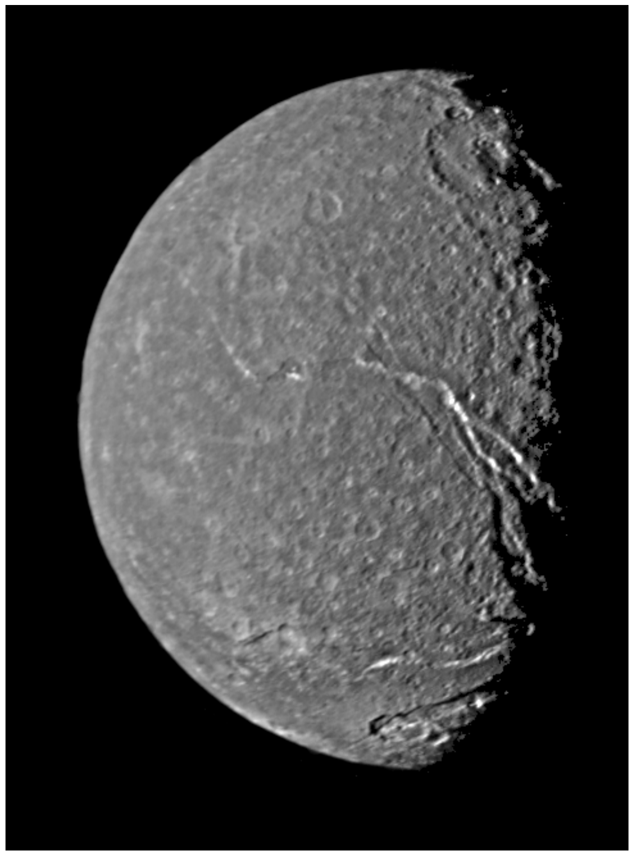

Figure 11.22: A view of Uranus' fourth intermediate moon, Titania. Like Ariel, this moon has moderate cratering (but lacks really large craters such as those seen on Umbriel), and also huge, large-scale faults, perhaps the result of surface expansion. (Courtesy of NASA.)

accretion and radioactive energy release is important only in relatively large moons such as the Galilean moons and Titan. For moons as small as those of Uranus, the amount of gravitational energy released is not enough to produce a large temperature rise, and radioactive heat produced by the rock component is quickly lost to the surface because of the small size. In the Uranian system the important sources of heating should be destructive impacts and tidal effects.

Consider tidal effects. All the intermediate-size moons of Uranus have orbital periods longer than the rotation period of the planet. The tidal bulges produced on Uranus by each of the moons will be dragged slightly ahead of the planet-moon line by the rotation of the planet. Thus the effects of tides on the moons from Miranda on out will be to increase the sizes of the orbits of the moon. (However, this effect depends on the mass of each moon as well as on it distance, so that the orbit of Ariel is actually increasing faster

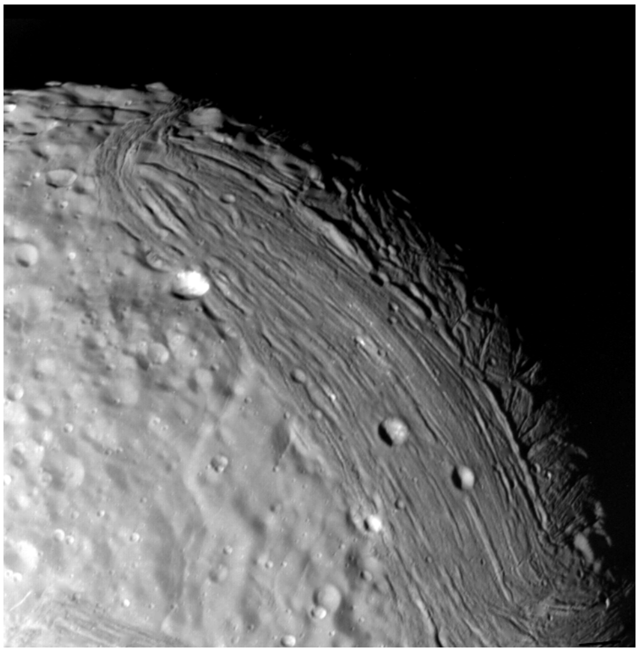

Figure 11.23: The smallest and innermost intermediate moon of Uranus is Miranda. This moon shows a relatively young, lightly cratered surface as well as very strange groove-and-ridge terrain that is not found anywhere else in the solar system. (Courtesy of NASA.)

than that of Miranda even though Miranda is closer to Uranus.) In the past, the moons have certainly had periods during which resonances between orbital periods existed (a past 3:1 resonance between Miranda and Umbriel is responsible for the present 4° inclination of Miranda's orbit to the equatorial plane of Uranus), but the computations of orbital evolution have not yielded any very convincing explanation of the heating of the three moons that have been resurfaced since the end of the period of heavy bombardment. The situation is particularly puzzling for Titania, an evolved moon located in between Umbriel and Oberon, the two moons with the least evolved surfaces. We do not have at present any secure explanation of the heating that led to resurfacing on three of the intermediate-size moons of Uranus.

A final question raised by the Uranus system would be to ask how both the planet and the accretion disk from which its moons system formed came to have such an unusual inclination to the plane of the planetary orbits around the sun. For the planet, this was probably the result of a near-catastrophic off-centre collision between the proto-planet and a second proto-planet of comparable size which drastically altered the direction of spin of the still forming body. However, there is no reason that further material accreting onto the planet would arrive in such a way as to orbit around the planet's highly inclined equator, and thus produce a moon system in this plane. How did the moon system end up forming in the planet's equatorial plane?

The effect that shifted the proto-planetary disk, and the moons that it formed, to the plane of Uranus' equator was again a tidal effect. Rapidly rotating Uranus bulges out at the equator. Because of this, the gravitational pull of the planet not only makes nearby moons and other material orbit around the planet, but any orbit not in the equatorial plane twists steadily around. This motion is rather like that of a spinning top whose axis is not quite vertical, so that the axis of the top moves around in a small cone about the contact point with the floor. This effect, acting on the orbits of moons, guarantees that orbits of different bodies orbiting out of the equatorial plane would sooner or later intersect. The various planetesimals would thus repeatedly collide until the debris settled into the present equatorial plane, where relatively long-lasting moons could finally form by (re-)accretion.

The strange case of Neptune's moons

The Neptune system is relatively poor in moons. It includes a single large moon, Triton, comparable in size to the smaller Galilean moons Io and Europa, as well as at least six small moons (all discovered by Voyager 2) orbiting close to the planet, four of which are even within the planet's Roche limit for ice-rich bodies. A single moon, Nereid, orbits in an extremely eccentric orbit well outside that of Triton.

Unlike the large moons of the other giant planets, the large moon Triton orbits around Neptune in the *opposite* sense to the planet's rotation, in an approximately circular orbit inclined to the plane of the planet's equator by about 23°. The moon's rotation about its own axis, however, is synchronized so as to keep one face always pointing towards Neptune. This extraordinary orbit makes it clear that Triton did not form in an accretion disk around Neptune, but was almost certainly captured more or less intact as it passed close to the planet. This would most likely have been possible as a result of a collision between Triton and a smaller moon already orbiting the planet, which could have slowed Triton down enough to insure that it would lack the energy to escape again from Neptune's gravitational grip.

As Triton gradually settled into a circular orbit from its initially very eccentric orbit (as a result of tidal effects), it would have disrupted any previous moon system, capturing some moons by collisions, and ejecting others from the system or causing them to crash into Neptune. The original satellite system was so greatly

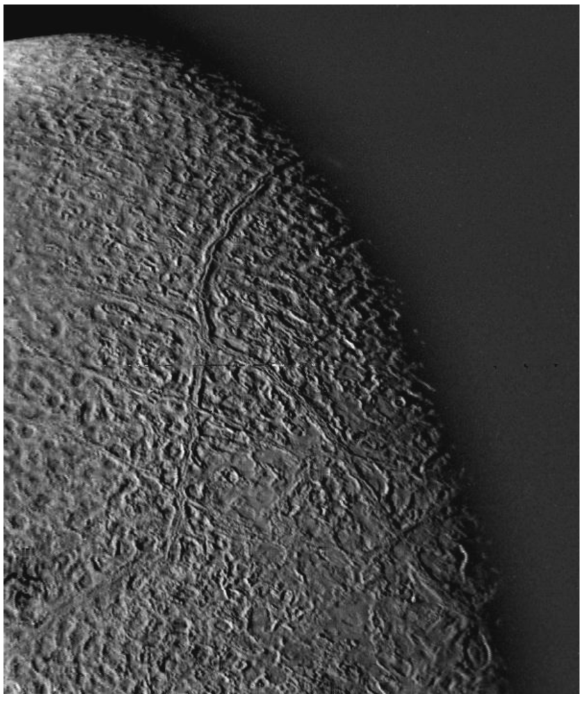

Figure 11.24: The trailing hemisphere of Neptune's single large moon, Triton, shows a vast system of roughly linear ridges that seem to have formed in plains already heavily deformed earlier. These features are probably the result of icy volcanism. Because of its appearance, this part of Triton's surface is known as "cantaloupe terrain". (Courtesy of NASA.)

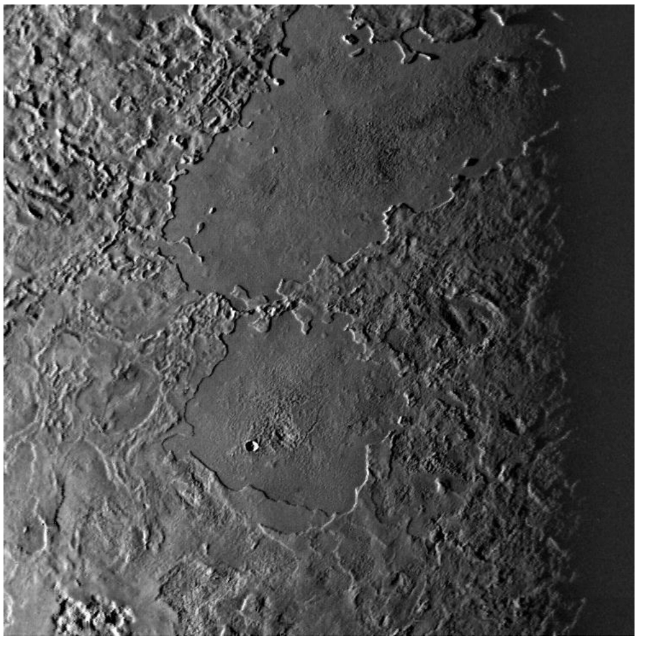

Figure 11.25: Another distinctive surface feature on Triton is several huge calderas or lakes that have been repeatedly filled with liquid water and drained. The lower lake in this image has a single small impact crater; generally, the surface of Triton is sufficiently recent that only a few small craters are found. (Courtesy of NASA.)

disturbed by Triton's arrival on the scene that all the remaining original satellites must have suffered destructive collisions with one another. The current set of small moons close to Neptune probably formed out of the debris of the earlier system, and even these moons have most likely been shattered by comet impacts since they re-formed, only to re-accrete again.

Exercise: Consider Triton shortly after it was captured into a very eccentric retrograde orbit around Neptune. Each time the moon passed close to the planet, it would raise tidal bulges on the near and far sides of the planet. Neptune's rotation would displace these bulges to lag behind the planet-moon line. With a sketch, explain how the near bulge would tend to slow Triton down a little in each close pass by Neptune, and how this would lead to gradual circularization of Triton's orbit.

With a mean density of 2050 kg m^{-3}, rock makes up more than 40% of Triton's mass. The surface of Triton is young, and there is no heavily cratered terrain. This is not surprising; the circularization of Triton's orbit by tidal effects after its capture must have deposited more than enough energy into the satellite to melt and differentiate it completely. The moon now consists of a core of rocky material surrounded by a mantle of ices. As a result of the intense tidal heating, the surface of Triton exhibits a variety of unique tectonic features. The winter polar region is covered with seasonal ice, and the surface there is not clearly visible. On the rest of the moon, the trailing hemisphere (relative to the moon's orbit about Neptune) is covered with terrain that looks like the skin of a cantaloupe: the ground is full of pits and dimples, criss-crossed by long, roughly straight ridges (Figure 11.24). The leading hemisphere is smoother, but has several frozen, terraced lakes, like volcanic calderas (Figure 11.25). In the southern hemisphere two powerful plumes, like geysers, rise to an altitude of about 8 km, where they form dense clouds which are stretched into a long tail by the winds.

Triton has a very thin atmosphere (about 10^{-5} as massive as that of Titan), mainly composed of N_2 together with a small amount of CH_4, but clouds and haze are seen, and this atmosphere is probably responsible for the transport of seasonal ices from one hemisphere to the other.

Little is known about the physical nature of the smaller moons except that they are quite dark in colour and so small that gravity cannot enforce spherical form; all are somewhat irregular in shape.

Exercise: We find planets and moons made primarily of rock, of a mixture of rock and ice, and of rock, ice, and gas. Which other combinations might be able to occur in nature? How might they form?

11.4 Planetary ring systems

Observations of the rings

As we have already seen, all four giant planets are now known to have ring systems. These ring systems exhibit a variety of complex phenomena that were hardly imagined before the 1970's, when the only known planetary rings were those of Saturn, and even for the rings of Saturn little detail can observed from the Earth. Theoretical astronomers have succeeded in explaining some of the observed features of these rings, but other aspects remain mysterious.

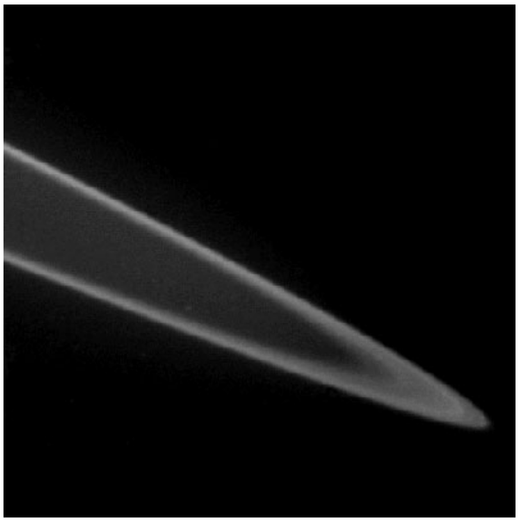

Figure 11.26: A view of Jupiter's ring from one of the Voyager probes, seen from behind the planet as the probe was leaving the Jupiter system. The main ring is visible as a distinct arc, while the wider dust halo is visible as a slightly brighter region inside the main arc compared to the dark sky outside. (Courtesy of NASA.)

The four ring systems are amazingly different from one another. Jupiter's single ring (Figure 11.26) is so faint that it can only be seen from behind the planet, and it has been observed only by the Voyagers and Galileo. It is made up of a single distinct main ring of rock and dust about 6000 km wide but only 30 km thick, orbiting Jupiter from about $1.72R_\mathrm{J}$ to $1.81R_\mathrm{J}$. It is known that the larger particles in the main ring are at least a cm in size, and are quite dark, like rock rather than water ice (note the rather unusual use of the word "particle", which in this setting can mean bodies even many meters in diameter). This ring is embedded in a larger and fainter halo of dust particles that extends inwards towards the planet. Two tiny satellites, Adrastea and Metis, orbit close to the outer edge of the main ring. The ring and these two moons are well within the planet's Roche limit, so each of the moons must be a solid block of rock.

Figure 11.27: This image of Saturn, taken as one of the Voyager spacecraft was leaving the giant planet, is somewhat overexposed for the planetary disk but clearly shows the complex structure of the major rings. In this image the clearly visible rings (starting closest to the planet) are the faint C ring, the bright B ring in which there seem to be many individual strands, the dark Cassini division, the featureless A ring with the faint dark Encke Division near its outer edge, and (as a faint thin line surrounding the A ring but separate from it) the narrow F ring. (Courtesy of NASA.)

Saturn's massive ring system, the only ring system easily seen from Earth, extends from about $0.1R_\mathrm{S}$ (6000 km) above the planet's cloud tops out to beyond the Roche limit. The outermost distinct ring is at about $2.32R_\mathrm{S}$, but a faint dusty ring extends out to about $8R_\mathrm{S}$. The main part of the ring system (the part that can easily be seen from Earth) is about 70 000 km wide. Even from the Earth it can easily be seen that Saturn's ring system is divided by a series of narrow dark rings into separate broad bands (Figure 11.27). These bands are designated by the letters (from the innermost to outermost) D, C, B, A, and F. Only the C, B and A rings are easily visible from the Earth; the D and F rings were discovered by the Voyager probes. Outside these are the very tenuous G ring and the wide, faint dusty E ring. The ring system from C to A is extremely thin compared to its width, only about 50 to 100 m from bottom to top, and when viewed edge-on from

Earth, it vanishes completely.

Although the rings seem fairly complicated when viewed from Earth, the images sent back by the Voyagers revealed a whole new level of complexity. What look like broad, almost featureless rings from Earth turn out to be composed of hundreds of thin, concentric strands. The broad rings also show wave-like distortions, and faint, wedge- or finger-like markings called spokes. The extremely narrow F ring shows kinks, warps, bright knots, and in places seems to be made of two or three thinner rings braided together.

Saturn's rings are quite massive compared to those of the other giant planets; with about 2×10^{19} kg of material in the C, B and A rings, there is nearly as much mass in the rings as in the innermost intermediate moon Mimas. Note that even from space probes we have never actually observed any of the ring particles in any planetary ring; we deduce the range of sizes and other properties of ring particles from the way in which they reflect sunlight, their thermal emission, and how well they reflect radar beams. From such data we deduce that Saturn's bright rings are mostly made of particles between about 1 cm and 5 m across, although there are some tiny dust grains present as well. The ring particles are highly reflective, and the spectrum of reflected light reveals the clear signature of water ice. The Roche limit for water ice lies within the A ring; just outside this are several tiny moons, including two moons whose orbits lie just inside and just outside the F ring.

Exercise: The rings of Saturn can easily be seen from Earth, while the rings of Jupiter and Neptune are not visible from Earth. Can you draw any conclusions from these facts?

The ring system of Uranus (Figure 11.28) is composed of some ten distinct narrow rings that orbit within wider dust bands. The rings are all well within the planetary Roche limit, and are bunched between $1.64R_{\mathrm{U}}$ and $2.00R_{\mathrm{U}}$. All but two of the distinct rings are amazingly narrow, ranging in width from about 1 up to 12 km. Two of the densest rings are wider; the η ring has a width of about 55 km, and the ε ring ranges from 20 to 96 km in width. All the rings seem to be extremely flat, with thicknesses perpendicular to their orbital planes of only a few tens of meters. A final remarkable feature is that most of the distinct rings are clearly slightly non-circular (i. e. they have slight eccentricity).

The outermost and densest ε (epsilon) ring includes particles ranging from dust grains up to boulders at least a meter in diameter, and probably contains most of the mass of the entire ring system. In the other rings the particles present include some dust grains as well as many fragments at least one cm in size. The particles in these rings are as black as soot, and reflect only a few percent of the light striking them. The rings of Uranus have only about 1/4000th of the mass of the rings of Saturn, but they contain considerably more mass than the rings of either Jupiter or Neptune.

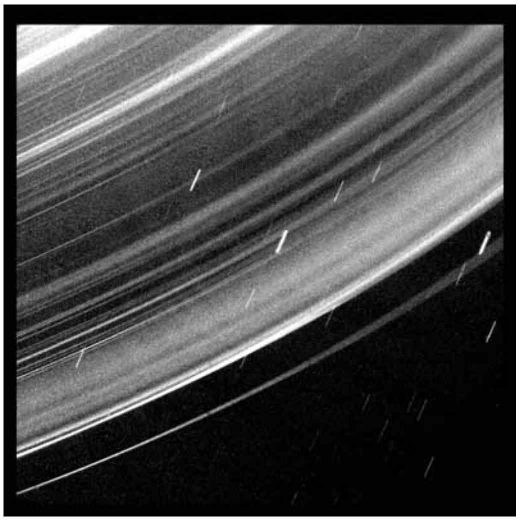

Figure 11.28: This view of the ring system of Uranus was obtained by Voyager 2 as it left Uranus behind on its way to Neptune. The part of the ring system imaged includes both the narrow rings that are visible from Earth and the wider dusty bands in which these are embedded. Some of the rings are clearly not symmetric about Uranus. The short oblique streaks are background stars whose apparent place in the sky changed during the time required to obtain this exposure. (Courtesy of NASA.)

The last of the giants to reveal its rings to Earth's inhabitants was Neptune, and again a ring system with many surprising and unique features was found. The rings were first discovered by occultations, and then extensively studied during the Voyager 2 passage through the Neptune system. Five rings are found, two distinct and narrow ones and three that are wider and fainter (Figure 11.29). The most remarkable feature of these rings is the fact that the outermost, narrow ring has several quite distinct bright clumps along part of its circumference, while the inner narrow ring is quite uniform along its length. Three of these clumps are visible in Figure 11.30.

Ring physics

The four giant planet ring systems have turned out to be far more complex and varied than anyone expected from the Earth-based views of Saturn's rings that were available before the start of space probe exploration of the outer solar system. The ring systems have a large

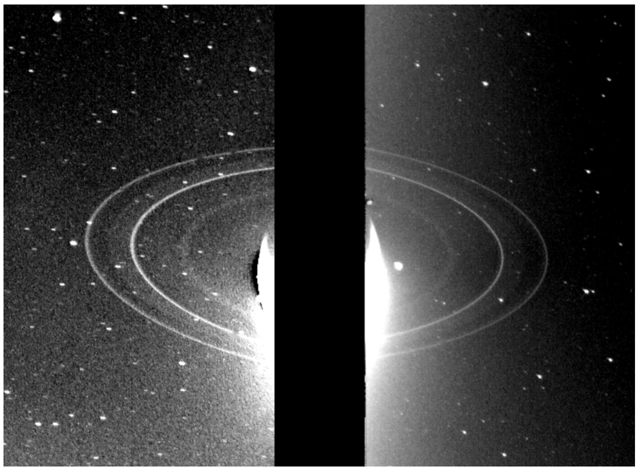

Figure 11.29: Two images of Neptune obtained by Voyager 2 clearly show the two bright,n arrow rings of Neptune. A thinner, more diffuse ring is faintly visible inside the inner sharp ring, and another in between the two distinct rings. (Courtesy of NASA.)

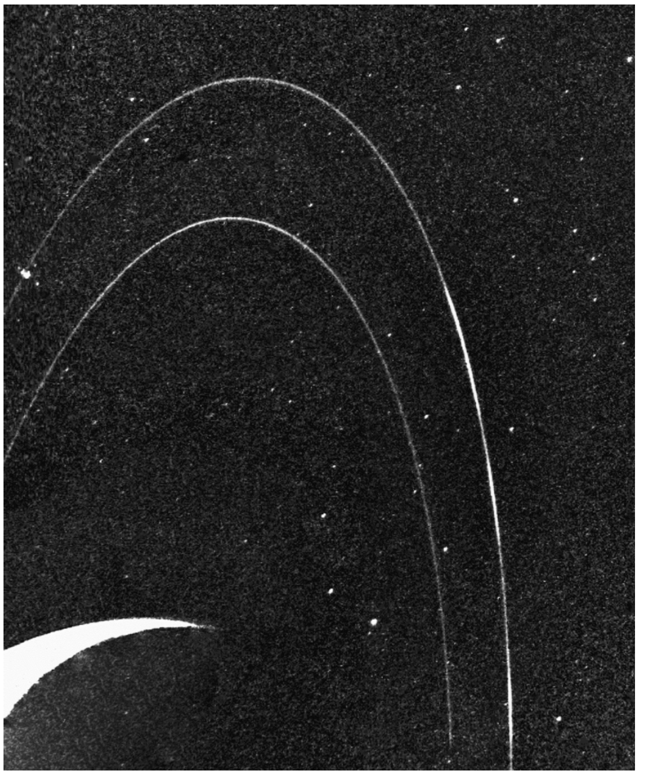

Figure 11.30: This image of Neptune's rings was obtained by Voyager 2 as it left the Neptune system behind (the back-lit disk of the planet is visible at the lower left). Only the two narrow distinct rings are visible. The inner is uniform along its length, but three quite distinct brighter segments are seen along the right side of the outer ring, where presumably some extra mass is concentrated. (Courtesy of NASA.)

range of masses, from the vast system of Saturn to the almost non-existent ring of Jupiter. All of the ring systems are extremely thin (typically tens or hundreds of m) through the ring plane, but individual rings range in radial width from 1 km up to the 70 000 km width of the Saturn's main rings. Most of the rings contain both relatively large particles (up to meters across in the rings of Saturn) and tiny dust grains. Many of the rings have distinct edges, and the wide B ring of Saturn seems to be made up of many thin concentric ringlets. The rings of Uranus are clearly eccentric. One of Neptune's rings has striking clumps, and Saturn's rings show "spokes". Mixed in with the rings are a number of tiny moons only a few km across. These remarkable systems have posed a large number of puzzles to planetary scientists, only some of which have found satisfactory answers up to now.

We can understand some of the phenomena of the ring systems if we recall the basic physical effects governing their orbital motions. The main force directing the motion of the particles that compose the rings is of course the gravity of the central planet, which forces the ring particles to move in essentially circular or elliptical orbits about the planet. A considerably weaker gravitational force is exerted on ring particles by the inner moons orbiting the planet, but as we will see, the weak moon forces can have an important effect on the shapes of rings. The ring particles also exert weak gravitational forces on one another. A third important force comes from the fact that the distinct rings have enough particles in them that collisions between ring particles are frequent, not rare, and this provides a means of changing the structure of a ring that is not available in systems with widely separated bodies, such as the main asteroid belt.

The faint, diffuse rings are mainly composed of dust particles which are subject to the same gravitational forces as the large particles from the central planet, nearby moons, and massive rings. However, in dust rings particle collisions are much rarer, and there are additional important forces such as the pressure of sunlight, effects due to thermal radiation, and even electrical and magnetic forces as these slightly electrically charged particles move through the magnetic field of their planet.

The completely dominant force of gravity from the central planet has the effect that all the ring particles move in orbits around the planet that are quite close to circles or ellipses. However, this fact alone is not enough to explain why rings are confined to a single (nearly equatorial) plane or why the various particle orbits are essentially similar (usually circular) in shape. These features of rings are consequences of the *collisions* between ring particles. As we have already mentioned in connection with the moons of Uranus, the equatorial bulge of the central planet (caused by its rapid rotation about it axis) causes the orbital plane of any ring particle whose orbit is tilted with respect to the equator of the planet to twist around slowly like a tilted spinning top. This effect guarantees that particles in such orbits will collide frequently with other particles in similar orbits. The final result is that the particles settle into orbits very close to the plane of the central planet's equator. Similarly, collisions between various ring particles moving in eccentric orbits will act like a kind of friction, and will cause the various orbits to gradually become circular. Thus, the facts that we find thin rings, very close to the planet's equatorial plane, and that the particles move in nearly circular orbits, are consequences of the occurrence of frequent collisions between ring particles.

Exercise: What disk-like collection of many particles in the solar system has the property that the particles collide only rarely, and are *not* confined to a very thin plane?

However, even when the orbits of ring particles are nearly circular and in the planet's equatorial plane, collisions between particles do not stop (although they become less frequent and less energetic). There are always particles whose orbits are a little different from those of their neighbors, and sooner or later these bodies run into others near them. Thus even when rings have become thin, flat, and circularized, there continues to be a kind of friction acting between particles orbiting the planet at different distances. This friction has the effect of speeding up the slower particles in the outer part of the ring where the orbits are slower, causing these particles to drift outwards into larger orbits. The faster particles in the inner part of a ring are slowed down slightly, causing their orbits to shrink slightly. The overall effect of the friction within a ring is to cause the ring to spread out slowly in width. Similarly, if a ring started out with a sharp inner or outer edge, we would expect this spreading effect to lead fairly rapidly to blurred edges.

Thus, the spreading effect of friction within a ring should lead planetary ring systems to be broad and diffuse. Planetary scientists were puzzled by the existence of gaps in Saturn rings, and even more astonished to find well-defined and extremely narrow rings around the other three planets. Clearly some additional processes must be acting to confine ring particles to distinct rings.

One such process is apparently the gravitational action of moons. Effects are produced by both the large moons outside the Roche limit, and also by smaller moons, usually only a few km across, that orbit close to the rings, often just inside or just outside a narrow ring. An example of such an effect is the outer edge of Saturn's B ring (the inner of the two brightest rings), where the Cassini division begins. Particles at this distance from Saturn have exactly half the orbital period of the innermost intermediate moon Mimas. (We say that these particles are in a 2:1 resonance with Mimas.) Particles orbiting here would encounter Mimas in the same place in their orbits every second revolution, so their orbits would quickly be forced to become quite eccentric, which would lead to collisions with other ring particles. This effect eliminates particles from this orbital radius, and causes the B ring to have a rather sharp edge.

Another kind of effect is "shepherding". Calculations have shown that a pair of tiny moons orbiting just inside, and just outside, a narrow ring are able to confine the particles of the ring, and prevent them from spreading out laterally. The tiny satellites Cordelia and Ophelia are apparently the shepherd moons for the ε ring of Uranus. A number of small moons were observed orbiting within the ring systems by the Voyager spacecraft, and many of the narrow rings may be confined by this mechanism. However, for other narrow rings, no shepherd moons have yet been discovered, perhaps because they were too small or dark to be detected by the Voyager cameras, or perhaps because shepherd moons are not the whole story.

There are at least tentative explanations for some of the other strange phenomena found in rings (eccentricity, clumpiness, waves, braiding, etc) but most of the explanations are rather difficult to understand in a simple way, and so we will not go farther into this problem.

Origin and evolution of ring systems

Perhaps the most interesting question about planetary rings is the question of their origin. Are the rings we see primordial, material that, because of its location inside the Roche limit, was unable to coalesce into moons, and is thus left over from the period when the planet and its large moons formed? A second possibility is that a ring results when a moon drifts inside the Roche limit (perhaps as a result of tidal interaction with the planet) and is torn apart by tidal forces. Similarly, a large comet passing close to the planet might be disrupted by tidal forces, and some of its debris could be captured. Still another possible origin would be that rings are the debris of small moons that are damaged or destroyed by high-speed collisions with passing comets.

Important clues about the origin of rings are found when we try to estimate how long the systems we observe could survive in essentially unchanged form. If we find that the observed ring systems should be able to persist largely unchanged for several Gyr, then a primordial origin is plausible. On the other hand, if the ring systems are expected to change greatly in millions of years or less, then a theory that creates ring systems continually or repeatedly may be more appropriate.

We have a couple of different ways of estimating the ages of the observed rings. For Saturn's bright, icy rings, one finds that over the age of the solar system, each particle should have collided with and accreted roughly its own mass in dark meteoritic dust. However, Saturn's ring particles are only as dark as would be expected after roughly 100 million yr of collisions. We conclude that these particles have *not* been sweeping up meteoritic dust for 4.5 Gyr: they are much younger than the solar system.

A second estimate of the ages of Saturn's rings comes from the tidal interactions between rings and their shepherd moons. The tidal interactions that allow these moons to shape the edges of rings also transfer rotation (angular momentum) between the rings and the moons. For example, the moon Prometheus (located just outside Saturn's A ring) would move the 2500 km that separates it from the A ring in only about 10 million yr. Thus the observed system will change substantially in some millions of years.

Other estimates of ring ages tend to agree that ring systems should change substantially, and perhaps even be completely disrupted or removed, within 10^7 or 10^8 yr. We conclude that the observed rings are probably relatively recent creations, not structures left over from the period of planet formation.

In this case, probably the most common method of making a ring is from the impact of a comet on a small moon. Comets passing close to one of the giant planets will typically be moving very rapidly; a comet passing within $2R_{\mathrm{S}}$ of Saturn will be moving at more than 25 km s^{-1}. If such a comet struck a moon with a diameter of a few tens of km, the impact could completely shatter both the moon and the comet. Calculations based on estimates of how many comets have passed close to each of the giant planets since their formation suggest that any moon close to the Roche limit of Jupiter or Saturn with a radius of less than 1–200 km is likely to have been disrupted at least once by such impacts. Thus the rings of Jupiter, Uranus and Neptune may well have been formed by such impacts.

The case of Saturn is a little more puzzling. Its ring system is so massive that it requires the recent disruption of an intermediate moon; a small one does not supply enough mass. This should have happened only a couple of times since Saturn formed, so the probability of finding the observed ring system at any particular time (e. g. now) is not high. It is of course possible that we have simply been lucky to have such an impressive ring system on hand now. Alternatively, perhaps a really large comet struck a small moon as it passed close to Saturn, disrupting the comet and enabling most of the debris to be captured, but this is also improbable.

The diffuse rings, composed almost entirely of tiny dust particles, are expected to have even shorter lifetimes than the main rings. The dust that makes up these very faint rings is rather easily driven out of the planet system by the steady orbit changes driven by radiation effects, or the particles are slowed by gas drag from the very thin outer atmosphere of the planet and spiral quickly in towards the planet. Thus such rings need almost constant replenishment. However, these rings are often found in orbits near one or another of the small inner moons, and it appears that bombardment of these small moons by fast meteoroids is frequent enough to replenish the dust rings with debris as fast as they are cleared away.

11.5 Pluto and Charon

The outermost planet, Pluto, is a body physically very different from the other planets, although perhaps not so different from some of the moons of the giant planets, or from other bodies in the Kuiper Belt. Furthermore, as the only planet not yet visited by any spacecraft for a close-up look, it remains in many ways uniquely mysterious.

The ninth planet was discovered only in 1930, by Clyde Tombaugh, an astronomer at Percival Lowell's observatory in Arizona. The discovery of the tiny body was the culmination of a long search for fifth massive outer planet, which had been hypothesized to explain what appeared to be irregularities in the orbits of Uranus and Neptune, in the same way that Neptune was discovered through its effects on the orbit of Uranus. (We now know that these orbital irregularities do not exist; no fifth giant planet is expected, or found.) With a position so far from the Sun that it requires 248 years to complete one orbit, it seemed quite appropriate to name the new planet after the Roman god of the underworld, Pluto, especially since the first two letters of the name are also Lowell's initials.

Pluto's orbit is much more eccentric ($e \approx 0.25$) than those of the other planets (except for Mercury), and quite strongly inclined ($i \approx 17°$) to the general orbital plane of the other planets. Because of the large eccentricity of its orbit, Pluto's perihelion ($q \approx 29.7$ AU) brings it closer to the Sun than Neptune for a short part of each revolution, while at its most distant retreat from the centre of the solar system ($Q \approx 49.5$ AU) it is near the outer edge of the Kuiper Belt. One consequence of this large eccentricity is that the solar energy falling on the planet varies by about a factor of three between perihelion and aphelion. The planet passed perihelion on 5 September 1989, and is now moving away from the Sun.

With an orbit that crosses that of Neptune, it would appear that Pluto would be in considerable danger of colliding with the larger planet. However, it is now clear that Pluto is locked in a 3:2 orbital resonance with Neptune (Neptune makes three revolutions about the Sun for two made by Pluto), so that Neptune passes Pluto only when the ninth planet is near the outermost part of its orbit, at a safe distance of about 17 AU from Neptune.

Exercise: Sketch the orbits of Neptune and Pluto and show how the 3:2 ratio of orbital periods leads to Neptune always passing Pluto when the smaller planet is in a particular part of its orbit.

Because Pluto is both small and very distant, for many years after its discovery almost nothing was learned about it except for its orbit. It appeared as an unresolved point of light (less than $1''$ in diameter) in ground-based telescopes, and so its diameter was essentially unknown, except for an upper limit (provided by a stellar occultation that did not occur) of 6500 km. No information at all was available about its mass. One discovery of note was made during this period: in the 1950's it was found that the planet's brightness varies regularly, by about 30%, in a 6.4 day cycle which is clearly the rotation period. This shows immediately that the planet has a surface which is not uniformly reflective.

The situation changed dramatically in 1978 when two astronomers at the U. S. Naval Observatory (only a few km from Lowell's observatory in Arizona) discovered that Pluto has a moon, so close to the planet (less than $1''$ away) that its light blends with that of Pluto except for observations under the best possible conditions. The new moon was quickly named Charon, after the boatman who ferrys the dead to the underworld in Greek mythology. This moon has made possible major advances in our knowledge of the tiny ninth planet.

Careful imaging from Earth, both from the ground and later from the Hubble Space Telescope (see Figure 11.31), showed that Charon moves around Pluto in an orbit with a semi-major axis of about 19,600 km. Furthermore, Charon's orbital period around Pluto is found to be synchronized with the planet's own rotation. The Pluto-Charon system is the only planet-moon pair in the solar system which is fully tidally locked, so that *both* bodies rotate synchronously with their mutual orbit, and each can be seen only from one side of the other. From the orbital period and semi-major axis, the mass of Pluto could be determined for the first time. The tiny planet has a mass of only about 0.0021 of the Earth's mass, less even than the mass of Neptune's moon Triton. (From the relative motions of Pluto and Charon on the sky, which has been successfully observed both from HST and from the ground, Charon's mass is found to be about 12% of that of Pluto.)

Figure 11.31: An image of Pluto and its moon Charon taken with the Hubble Space Telescope when the two bodies were near their maximum apparent separation on the sky. (Courtesy of NASA.)

The plane of Charon's orbit, and of Pluto's equator, is oriented roughly perpendicular to the plane of the solar system, and at present the rotation axes of the two bodies are roughly parallel to the direction of their orbital motion around the Sun. As a result, we had the good luck that only a few years after the discovery of Charon, the Earth moved into Charon's orbital plane, and the planet and moon began to eclipse one another. Timing of these eclipses ("mutual events") made possible, for the first time, accurate measurements of the radii of the two bodies, which were later confirmed by direct imaging by the HST. It is found that the radius of Charon (about 590 km) is about half that of Pluto itself (about 1150 km), making Charon the largest moon, relative to its planet, in the solar system. Combining the planetary radius with the mass, we find that Pluto has a mean density which is about 2050 kg m^{-3}. The planet therefore has a composition, assuming as usual a mix of rock and ice, which is around 60 – 75% rock. This proportion is one of the largest values found among the moons of Saturn, Uranus, and Neptune, which generally have between 50 and 60% rock; only Titan, Oberon, and Triton have similarly large rock fractions. Charon is found to have a similarly large mean density, about 1800 kg m^{-3}, so its proportion of rock is similar to that of Pluto.

One final discovery of note was that of strong evidence in the infra-red spectrum of the presence of CH_4 frost on the planet's surface. Because this ice has an appreciable vapour pressure even at the very low temperature expected on Pluto's surface (roughly 50 K), it was realized that the planet must have a very thin atmosphere. This was confirmed in 1988 when the occultation of a bright star occurred with the kind of gradual dimming expected due to this atmosphere. In contrast, Charon shows no trace of CH_4, but does appear to have water ice at its surface.

Because Pluto has never been visited by a spacecraft, we have no close-up pictures of its surface, and thus no direct evidence as to whether the planet is differentiated or homogeneous. However, there are several reasons to think that Pluto probably is differentiated. Recall that, like the moons of the giants, differentiation occurs if the *ices* are melted, which only requires a temperature rise of about 200 K from the current temperature near 50 K. Some of the arguments pointing towards differentiation are the following. (a) Accretion of the planet could have provided enough energy to melt the ices if most of the gravitational energy released was retained by the planet. (b) With the high rock fraction in the planet, radioactivity would probably have released enough energy by now to cause melting. (c) The observed surface CH_4 frost would not be present if the CH_4 were mixed with other, less volatile ices such as water ice; the surface methane would have evaporated and escaped from the planet, leaving a surface of other, less volatile ices. We deduce that the surface coating of CH_4 may be at least some km thick, which appears to require that differentiation has taken place. (d) Finally (see below), we think that Charon became bound to Pluto after a giant impact somewhat like that now thought to have formed the Earth's Moon; this would also probably have released enough energy to melt the planet.

If we are correct that Pluto has differentiated, a plausible interior model would have a core of rock roughly 900 km in radius, with a layer of mostly water ice above, probably about 250 km thick, that extends to within some km of the surface. Above all is a thin layer, some km thick, of methane ice. The internal temperature could be somewhere in the vicinity of the melting point of water, although the surface is at a chilly 40 to 50 K.

The origin of an apparently unique, tiny planet in the outer solar system has long been a mystery. Pluto's Neptune-crossing orbit led to the proposal that Pluto might once have been a moon of Neptune that suffered a close encounter with Triton, ejecting Pluto from the Neptune system and putting Triton into its present unusual retrograde orbit. This idea now seems very unlikely for several reasons: first, because Pluto is less massive than Triton, Pluto would not be able to reverse Triton's orbit; secondly, because Pluto would leave Neptune in an orbit which would intersect that of Neptune, the two planets would rather quickly collide; and finally, it is not at all obvious how Pluto could get from a Neptune-intersecting orbit into its present strongly Neptune-avoiding 3:2 resonance orbit.

Most scientists now suppose that Pluto formed as an independent body, by accretion of planetesimals, early in the history of the solar system. The recent discovery of hundreds of bodies in the Kuiper Belt, many of which have dimensions approaching those of Charon, certainly makes this origin seem quite plausible. Furthermore, the existence of the many bodies in the Kuiper Belt may offer a reasonable explanation of the origin of the Pluto-Charon system. With probably thousands of bodies in this region, it does not seem too unlikely that Pluto might have suffered a collision with a large one, leading to formation a moon out of the collision debris. Note, however, that because of the small size of the bodies involved, such a collision would have occurred with only about 1/10th the relative velocity that the Earth and its moon had at the time of their collision, so the heating effect – though probably enough to melt Pluto – would have been much less catastrophic. Nevertheless, enough volatiles might have been lost from the Pluto-Charon system to account for its relatively large mean density, and for that of Charon.

Although the solar system's outermost planet is still known quite imperfectly, it is no longer the total mystery it was for the first half century after its discovery. Instead, its particular nature – unique, but clearly related to other objects we have gotten to know – reminds us again of the immense variety and strangeness of the small system of bodies that travel through the Milky Way together with our Sun.

11.6 Mathematical aspects

Gravity, hydrostatic equilibrium, and cooling

Many of the ideas that we have discussed in the Mathematical Aspects sections of earlier chapters are relevant to the giant planets and their moons. For example, we have discussed in Chapters 8 and 9 how the mean density of a planet or moon may be used with reasonable guesses about possible substances making up the body to estimate the relative fractions of these components, and if they are separated, to estimate the size of the core and the mantle.

Exercise: Data about Jupiter's moon Callisto may be found in Table A.3. Assume that this moon is composed of partly of ice (of density 900 kg m^{-3}) and partly of rock (of density 3500 kg m^{-3}). (a) Show that the general expression relating the mass fractions f_1 and $f_2 = 1 - f_1$ of two composition components of densities ρ_1 and ρ_2 to the observed mean density $\overline{\rho}$ of the moon is simply

$$\overline{\rho} = f_1\rho_1 + f_2\rho_2. \tag{11.1}$$

(b) Use the measured mean density of Callisto to estimate the fraction of the total mass that is rock, and the fraction that is ice. Do you find that these components each make up roughly half of the moon's mass?

The estimate that you derived earlier (Equation 6.8) for the maximum possible temperature increase that could be produced by the conversion of gravitational accretion energy into internal energy should apply to the accretion of the Galilean moons.

Exercise: Assume that the material from which Europa formed had an initial temperature of 100 K. Could accretion have supplied a large enough temperature increase to melt ice (273 K) and lead to separation of a rocky core from an ice mantle?

We have also seen that internally and in any atmospheres present, the material of these bodies will be approximately in hydrostatic equilibrium, which we examined in Chapters 3 and 8. This fact may be used to estimate conditions both in the deep interior of a planet or moon, and in its atmosphere.

Exercise: Use the reasoning discussed in Chapter 3 to estimate the order of magnitude of the pressure at the centre of Saturn. Does your result agree reasonably with the value you infer from Figure 11.5?

Exercise: At the top of Jupiter's uppermost cloud deck, the pressure is about 6×10^4 Pa. Assuming for simplicity that the gas is all H_2 molecules, estimate the local number density of molecules (number per m^3), and the total mass of gas per m^2 above this level. (You may want to look back at Chapters 3 or 10.)

Another aspect of the giants and their moons that can be studied with physics we have already met is the surface temperatures expected from equilibrium with sunlight. Recall from Equation 7.3 that this depends on the albedos for visible and infrared radiation, as well as the distance of the bodies from the sun.

Exercise: Europa reflects 58% of the visible light falling on it, while Callisto reflects only 13%. Assuming that both moons reflect only about 5% of infrared radiation, determine the average surface temperatures of the two moons. Why is the surface of Europa cooler than that of Callisto? (Hint: don't just directly apply Equation (7.4).)

Growth of a surface ice layer

Yet another problem we can examine is the rate of growth of the ice layer over a sea of water on a moon of one of the giant planets shortly after it first forms. Let's assume that the layer is so thin that the ice layer can be treated as flat, not curved around the moon. We also assume that the only way in which heat is carried out through the icy crust is by thermal conduction; this is the key assumption, because if solid-state convection is able to start in the ice, heat will be transported much more efficiently than by simple conduction. Finally, we assume that the surface layer of the ice is held at a constant temperature $T_{\rm s}$ by the balance between incoming sunlight and thermal re-radiation [Equation (7.4)], and that the bottom of the ice layer is held at the melting temperature of the water in the sea, $T_{\rm m}$ (why?).

Now a certain amount of energy will leak out from the sea below to the surface of the ice sheet. We measure this heat leakage by the heat flux q, the energy carried through the ice sheet per m^2 per s. From Equation (6.9), at a time t when the ice sheet has reached a thickness $D(t)$, this heat flux is given by

$$q = k_{\rm c}\frac{T_{\rm m} - T_{\rm s}}{D}. \tag{11.2}$$

This means that in some time dt, the amount of energy lost from one square m of the sea is $q\,dt$. Now let us suppose that the rate of heating of the sea by radioactive energy release in the core is insignificant compared to the rate of heat leakage through the ice layer, which it will be shortly after the liquid water at the surface of the moon starts to freeze. In this case, the heat lost from one m^2 at the top of the sea will have to be provided by freezing a thin layer of ice onto the bottom of the ice sheet; the energy needed will be the latent heat of fusion released by this process. The amount of latent heat released in freezing one m^3 of ice is $L\rho$, where L is the latent heat per kg and ρ is the density of the ice. Thus to provide an amount of energy $q\,dt$ from one m^2 at the top of the sea, a volume $dV = 1 \times dD$ must freeze such that

$$q\,dt = L\rho\,dV = L\rho\,dD \tag{11.3}$$

where dD is the thickness by which the ice layer increases in dt. Then

$$L\rho\frac{dD}{dt} = k_{\rm c}\frac{T_{\rm m} - T_{\rm s}}{D}. \tag{11.4}$$

This equation may be rewritten as

$$D\frac{dD}{dt} = \frac{k_{\rm c}(T_{\rm m} - T_{\rm s})}{L\rho} \tag{11.5}$$

which is easily integrated (with $D = 0$ at $t = 0$) to yield

$$D(t) = \left(\frac{2k_c(T_m - T_s)}{L\rho} t \right)^{1/2}. \quad (11.6)$$

Thus we see that the thickness of the ice layer grows as the square root of the time elapsed since the start of freezing, rather than linearly with time. (This is the basic reason that most lakes on Earth do not freeze all the way to the bottom during the winter.)

Exercise: Using data from Tables 2.4 and 6.7, estimate the amount of time that would be required for a surface ice layer to reach a thickness of 20 km on Jupiter's moon Ganymede.

Tidal disruption

We next turn to the subject of tidal disruption. Unlike the description of the orbital movement of a small body around a larger one, tidal disruption cannot easily be described exactly. Instead, we look for a way to get an estimate of the size of this effect by suitable approximations.

To treat tidal disruption approximately, recall that the force of gravity exerted by one point mass (or spherical body) of mass M, radius R, and density ρ_M, on another of mass m, radius r, and density ρ_{rmm}, is

$$F = GMm/a^2, \quad (11.7)$$

where G is the gravitational constant and a is the separation of M from m. Now to estimate the tidal effect of M on m, we divide m mentally into two halves, one on the side near M and the other on the side opposite M. Suppose now that the body m is held together by gravity, so that the force holding these two halves together is only gravitational; then the attractive force of gravity F_g of the near half of m on the far half of m is (rather approximately)

$$F_g \approx \frac{G(m/2)(m/2)}{(r)^2} = \frac{Gm^2}{4r^2}. \quad (11.8)$$

Opposing the self-gravity of m is the tidal effect of M. The attraction of M on the near half of m is approximately

$$\begin{aligned} F_n &\approx \frac{GM(m/2)}{(a - r/2)^2} = \frac{GMm}{2a^2(1 - r/2a)^2} \\ &\approx \frac{GMm}{2a^2}(1 + r/a), \end{aligned} \quad (11.9)$$

where have used a first-order Taylor expansion of the parenthesis in the denominator to get the last approximate equality. Similarly, the gravitational force due to M on the far side of m is approximately

$$F_f \approx \frac{GM(m/2)}{(a + r/2)^2} \approx \frac{GMm}{2a^2}(1 - r/a). \quad (11.10)$$

Because of the larger distance of the far half of m from M compared to the separation of the near half of m and M, F_n is somewhat larger than F_f. Thus M will tend to attract the near half of m toward itself more strongly than the far half. This is the origin of the tidal effect. To keep the two halves of m moving together, we need the attractive force F_a of the near half of m on the far half, and the attractive force by the far half of m on the near half to be strong enough that the net force on each half of m equal, so that they follow the same orbits. This requires that $F_f + F_a = F_n - F_a$, or

$$\begin{aligned} F_a &\approx (F_n - F_f)/2 \\ &\approx \frac{GMm}{4a^2}[(1 + r/a) - (1 - r/a)] \\ &= GMmr/2a^3. \end{aligned} \quad (11.11)$$

Now for the small body to be stable against gravitational disruption, its self-gravity must be large enough to supply the required F_a, so stability occurs for

$$F_g \approx Gm^2/4r^2 \geq F_a \approx GMmr/2a^3, \quad (11.12)$$

or

$$a > r(2M/m)^{1/3} \approx r_{\text{Roche}}. \quad (11.13)$$

But now the density of M is just $\rho_M = M/(4\pi R^3/3)$, while $\rho_m = m/(4\pi r^3/3)$, so we may rewrite r_{Roche} as

$$\begin{aligned} r_{\text{Roche}} &\approx r(2M/m)^{1/3} \\ &\approx R(2\rho_M/\rho_m)^{1/3} \\ &\approx 1.26R(\rho_M/\rho_m)^{1/3}. \end{aligned} \quad (11.14)$$

From this form it is easy to see that tidal disruption only occurs for separation a not much greater than the radius R of the larger body, and that larger density ρ_m decreases r_{Roche} and promotes stability of m.

Although the derivation above is quite rough, the result shows the correct dependence of r_{Roche} on the two densities, with a coefficient which is of the right order of magnitude. The exact result for two fluid bodies of uniform density is

$$r_{\text{Roche}} = 2.45R(\rho_M/\rho_m)^{1/3}. \quad (11.15)$$

This is the value of the Roche limit that should be used for computations.

Note that a single object held together by internal forces that are stronger than gravity (a solid piece of ice, for example) will not be disrupted even if it ventures inside the Roche limit of a larger body.

Exercise: Estimate the Roche limit [using Equation (11.15)], expressed in units of planetary radii, for tidal disruption of (a) the Earth's Moon by Earth and (b) Io by Jupiter. (c) Is the outer edge of Jupiter's ring, at $1.81R_{\mathrm{J}}$, inside or outside the Roche limit for a solid body made of water ice that orbits Jupiter?

11.7 References

Atreya, S. K. et al. 1999, "A comparison of the atmospheres of Jupiter and Saturn: deep atmospheric composition, cloud structure, vertical mixing, and origin", *Plan. Space Sci.*, 47, 1243. Current problems in understanding the atmospheres of Jupiter and Saturn are summarized in detail in this review article.

Binzel, R. 1990, "Pluto", *Sci. Am.*, Jun., 50. Binzel provides a wide-ranging discussion of what was known about Pluto, including first attempts to map its spotted surface, before HST observations began.

Burns, J. A., Hamilton, D. P. and Showalter, M. R. 2002, "Bejeweled Worlds", *Sci. Am.*, Feb., 64. A nicely illustrated view of the rings of the four large outer planets.

Cuzzi, J. N. and Esposito, L. W. 1987, "The rings of Uranus", *Sci. Am.*, Jul., 52. Planetary rings are much more complicated, and probably ephemeral, than we ever imagined. This article discusses some of the puzzles they offer.

Drake, S. and Kowal, C. T. 1980, "Galileo's sighting of Neptune", *Sci. Am.*, Dec., p. 74. Neptune was observed long before it was discovered.

Esposito, L. W. 1993, "Understanding Planetary Rings", *Ann. Rev. Earth Planet. Sci.*, 21, 487. Esposito discusses many of the main problems encountered in understanding planetary rings, and describes some of the answers that have been proposed, in the somewhat technical (but largely nonmathematical) review.

Grasset, O., Sotin, C. and Deschamps, F. 2000, "On the internal structure and dynamics of Titan", *Planet. Space Sci.*, 48, 617. This somewhat technical article describes how the interior of Titan (as well as Ganymede and Callisto) may have evolved since the moon was formed.

Hubbard, W. B. 1984, *Planetary Interiors* (New York: Van Nostrand Reinhold). Chapter 8 of this excellent advanced textbook has a very clear discussion of how we deduce the interior composition and structure of the giant planets and their moons.

Ingersoll, A. P. 1981, "Jupiter and Saturn", *Sci. Am.*, Dec., 90. Ingersoll describes efforts to understand both the interior structure and atmospheres of the two giants in the light of Voyager data.

Ingersoll, A. P. 1987, "Uranus", *Sci. Am.*, Jan., 38. Deductions from the Voyager 2 fly-by of Uranus in 1986 are surveyed.

Johnson, T. V. 1995, "The Galileo mission", *Sci. Am.*, Dec., 44. The most ambitious space probe mission to Jupiter is described, mainly focusing on the technical challenges.

Johnson, T. V. 2000, "The Galileo mission to Jupiter and its moons", *Sci. Am.*, Feb., 40. In spite of major technical problems, the Galileo space probe has made possible a detailed study of Jupiter and its large moons.

Johnson, T. V. et al 1987, "The moons of Uranus", *Sci. Am.*, Apr., 48. Voyager 2 not only discovered 10 new moons of Uranus, but also showed that three of the larger ones have had a geologically active past.

Laeser, R. P. et al 1986, "Engineering Voyager 2's encounter with Uranus", *Sci. Am.*, Nov., 36. A lot of in-flight engineering of Voyager 2 made possible its spectacularly successful fly-by of Uranus.

Malhotra, R. 1999, "Migrating planets", *Sci. Am.*, Sep., 56. Calculations suggest that the orbits of Neptune, Pluto, and some of the bodies of the Kuiper Belt have changed substantially since the solar system formed.

Miner, E. D. 1990, "Voyager 2's encounter with the gas giants", *Phys. Today*, Jul., 40. Miner describes the many technical challenges of the Voyager missions to the outer planets, and summarizes the main scientific results.

Owen, T. C. 2000, "On the origin of Titan's atmosphere", *Planet. Space Sci.*, 48, 747. Owen explains the arguments leading to the present ideas about how Titan's atmosphere originated.

Pappalardo, R. T. et al. 1999, "The hidden ocean of Europa", *Sci. Am.*, Oct., 54. Deep under its icy crust, Europa may have an ocean of liquid water kept warm by tidal heating.

Peale, S. J. 1999, "Origin and evolution of the natural satellites", in *Ann. Rev. Astr. Ap.*, 37, 533. Peale presents a comprehensive (but fairly advanced) discussion of all the satellite systems in the solar system with particular emphasis on how tidal effects have changed both orbits and in some cases internal structure.

Pollack, James B. 1984, "Origin and history of the outer planets: theoretical models and observational constraints", in *Ann. Rev. Astr. Ap.*, 22, 389. This review article discusses theories of the formation of the outer planets and their moons, and how observational data can help to select among competing theories.

Pollack, J. B. and Cuzzi, J. N. 1981, "Rings in the solar system", *Sci. Am.*, Nov., 104. The Voyagers revealed much unexpected complexity in the rings of the outer planets, especially those of Saturn.

Showman, A. P. and Malhotra, R. 1999, "The Galilean satellites", *Science*, 286, 77. This article reviews the many kinds of new information gathered about the four large moons of Jupiter by the Galileo orbiter, and has a lot of useful references to more detailed articles.

Soderblom, L. A. and Johnson, T. V. 1982, "The moons of Saturn", *Sci. Am.*, Jan., 101. The two Voyagers have revealed that most of the moons of Saturn are heavily cratered, and Enceladus shows evidence of complex evolution.

Stern, S. A. 1992, "The Pluto-Charon system", *Ann. Rev. Astron. Astrophys.*, 30, 185. Stern examines the structure and development of both Pluto and its moon in this comprehensive review.

Stevenson, D. J. 1982, "Interiors of the giant planets", in *Ann. Rev. Earth Planet. Sci.*, 10, 257. A clear review describing how we obtain information from a combination of theory and observation about the internal structure of the giant planets.

11.8 Review questions

11.1 How can we determine the mass and radius of a giant planet? Using densities derived from these data, does it appear that all four giant planets have similar chemical composition?

11.2 Why is it useful to simplify the composition of the giants to gas, ice, and rock?

11.3 How is a "model" of a giant planet made? What is the difference between a "cold" model and a "warm" one?

11.4 What are the main ways in which Uranus and Neptune differ from Jupiter and Saturn?

11.5 What processes might have formed the giant planets? What data and arguments allow us to choose one possibility over others?

11.6 How hot are the giants inside? Where does internal heat come from, and how does it reach the surface?

11.7 What can we deduce about the chemical composition of the moons of the giant planets?

11.8 Why does water ice predominate over frozen methane, ammonia, or carbon dioxide in the moons of the giants?

11.9 What heat sources may have powered tectonic activity in the moons of the giant planets?

11.10 How do surface features help us to understand the history of the moons of the giant planets? How can we tell if a moon is differentiated or not?

11.11 How could a moon be captured by a giant planet?

11.12 What processes change planetary rings with time, and cause them to evolve?

11.13 How can shepherd moons exist inside a planetary Roche limit?

11.14 How may rings have formed? Are they probably permanent features of their planets?

11.9 Problems

11.1 Jupiter has about 5×10^{29} molecules of H_2 and 1×10^{29} molecules of He per m^2 above the NH_3 ice cloud tops, where the local temperature is $T_c = 140$ K. (Note: these are *total* numbers of molecules above each square meter, *not* local number densities.) Write down an expression for the pressure p_c at the cloud tops using this information, and evaluate it. (b) Assume that the atmosphere is convective below the NH_3 cloud tops, and that the temperature increases with depth s at the constant adiabatic lapse rate of $dT/ds = 2 \times 10^{-3}$ K m^{-1}. Write down an expression for the temperature $T(s)$ below the cloud tops, and then the equation of hydrostatic equilibrium including explicitly the variation of temperature with depth. (c) Integrate the resulting equation of hydrostatic equilibrium to find the variation of $p(s)$ with depth. Use the known values of pressure and temperature at the NH_3 cloud tops to evaluate any constants. (d) The bottom of the water ice clouds occur at a depth where the temperature is about 280 K.

Evaluate the depth s_w of this level below the level of p_c, and find the pressure $p(s_w)$ there.

11.2 At the top of the NH_3 clouds in Saturn's atmosphere, the temperature is about 110 K and the pressure is about 0.5 bar. Below this level the atmosphere is convecting and the temperature increases with depth s at a constant rate of $dT/ds = 7 \times 10^{-4}$ K m^{-1}. (a) Write down an explicit expression for the value of temperature as a function of depth, $T(s)$, and the expression for the equation of hydrostatic equilibrium including this explicit variation of $T(s)$. (Eliminate density from the equation of hydrostatic equilibrium using the ideal gas law, assuming that Saturn's atmosphere is composed of pure H_2.) (b) Integrate your equation of hydrostatic equilibrium to find an explicit expression for the pressure $p(s)$ as a function of depth below the NH_3 cloud tops. Use the data above to evaluate any constants. (c) Use the ideal gas law and your expressions for $p(s)$ and $T(s)$ to find an expression for the density $\rho(s)$ as a function of depth. (d) The separation of the two nuclei in H_2 is about 7.4×10^{-11} m, so two molecules will essentially touch when their separation is of the order of 10^{-10} m. When the density is high enough for the mean separation between molecules to be of this order, H_2 will surely be essentially liquid rather than gaseous. Imagine for counting purposes that the H_2 molecules in a liquid are arranged in a simple lattice of rows, columns, and layers with all the spacings between adjacent molecules being 1×10^{-10} m. What is the density of H_2 in this state, in kg m^{-3}? (e) Use your results from parts (c) and (d) to estimate the depth below the cloud tops at which Saturn's interior changes from a gas to a molecular liquid.

11.3 Atoms of all atomic numbers have radii comparable to twice the Bohr radius $a_o = 5.29 \times 10^{-11}$ m. In general, one cannot crowd atoms significantly closer together than a mean spacing of $\sim 4a_o$ without providing enough energy to disrupt the electrostatic structure of the atoms (i.e. without providing enough energy to detach at least the valence electrons, thus ionizing the atoms). (a) Calculate the approximate density of liquid hydrogen, of a common rock such as Mg_2SiO_4, and of lead, assuming that in all cases the spacing between atoms is approximately $4a_o$. (Recall that Mg_2SiO_4 has seven atoms, not one.) Compare your estimates to measured values. Considering that this is only an order-of-magnitude estimate, do you think that the assumption of a universal atomic size is roughly correct? (b) Assume that a giant planet of radius R is composed entirely of H, with a (compressed) density of 300 kg m^{-3}. Compute the total number of atoms in the planet, and the gravitational energy released in forming the planet. How large must the radius R be in order for the gravitational energy to be just large enough to provide enough energy to remove the electron from each H atom, assuming that this requires of order $E_{\rm coulomb} \sim e^2/4\pi\varepsilon_o a_o$ Joules per atom? (This is a rough estimate of the minimum mass a planet must have to convert all molecular H into metallic H.) Compare your calculated mass to that of Jupiter.

11.4 (a) Use the measured mean density of Ganymede from Table A.3 to determine the mass fractions of ice (density 900 kg m^{-3}) and rock (density 3200 kg m^{-3}). (b) Assume that Ganymede has separated into a core-mantle structure. Using the mass fractions from (a), determine the radius of the core and the thickness of the mantle. (c) Estimates of the densities of each composition component are always uncertain because of imprecise knowledge of how the moon accreted. Assume that the density of ice is relatively certain, but that the density of the rocky component is uncertain by ± 400 kg m^{-3}. Estimate the corresponding uncertainty in the radius of the core.

11.5 This is problem to examine the break-up of Comet Shoemaker-Levy 9 (SL-9) which struck Jupiter in 1994 (see Figure 7.11). Comet specialists have concluded that SL-9 passed Jupiter in July 1992 at a distance of about 113,000 km from the planet's centre, well within the Roche limit, when the comet broke into a number of fragments. In this problem, we will see that the break-up in 1992 explains very nicely the fact that the pieces arrived at Jupiter in July 1994 over a period of about 1 week, rather than all together. (a) Using the time between the last two approaches to Jupiter and Kepler's third law, determine the semi-major axis of the last orbit of SL-9 around Jupiter. (Be sure to use Jupiter's mass, not that of the Sun!) (b) Now let us assume that the comet was about 5 km across when it broke up, and that it broke up at its closest approach to the planet. Thus the *nearest* pieces of comet to Jupiter were at a distance r_1 about 5 km less than the distance r_2 of the *farthest* pieces. However, at the moment of breakup, all the pieces were travelling at the *same* velocity. Use the *vis viva* equation (Equation 1.14), and the fact that the velocities of all fragments at break-up were the same, to determine the difference between the semi-major axes of the orbits of the "extreme" fragments, 2.5 km closer and farther from Jupiter than the average. (This can be done directly on your pocket calculator if you keep enough decimals, or more elegantly using Taylor expansions.) Then use Kepler's third law again to find the difference in the extreme periods. You should find that the difference in the orbital periods of the fastest and slowest fragments is of the order of some days, neatly explaining the spread in arrival times of the pieces of SL-9.

11.6 One of the major factors in the evolution of the small moons close to Saturn has probably been impacts by comets. In this problem we look at just how destructive a comet impact could be. The problem is to estimate the relative velocity of a plausible comet with a small moon when they collide. (a) Consider a spherical comet composed of water ice and having a radius of 10 km, entering the inner solar system from an aphelion distance of 1×10^4 AU. What is the velocity of this comet as it reaches Saturn's distance from the Sun, 9.54 AU? (b) Now calculate approximately the effect of entering the gravity field of Saturn by using conservation of energy, starting with the velocity calculated in part (a), which we can assume applies at, say, 5×10^6 km from Saturn, to determine the velocity of the comet as it reaches a distance from Saturn of 150,000 km (about the orbital radius of Janus or Epimetheus). (c) Finally, suppose that the comet strikes a small moon at this distance from Saturn. Suppose, to maximize the effect, that the moon is travelling around Saturn in a circular orbit and encounters the comet head–on. What is the approximate relative velocity $v_{\rm rel}$ at the moment of impact? What is the kinetic energy release in the impact, computed as $m_{\rm comet} v_{\rm rel}^2/2$? (d) Consider a moon with a uniform mixture of 0.75 ice and 0.25 rock, of radius $R_{\rm m}$. What is the minimum radius the moon must have for the collision to bring in less energy to the system than the gravitational binding energy of the moon, and thus be unable to completely disrupt the moon?

11.7 As discussed in the text, Titan may have an icy crust with a sea beneath. One way in which this could happen would be if the ices that formed Titan contained a few percent of NH_3, which seems likely. The ammonia would act as an anti-freeze, keeping water liquid down to a temperature of about 170 K. In this case, the cold liquid at the bottom of the icy surface crust would keep the ice so cold that solid state convection would be very inefficient, and most of the heat loss through the crust would be by simple thermal conduction. Assuming that this is the case, we can estimate the thickness of the ice crust on the moon. (a) First, we need an estimate of the heat flowing out of the rocky core of the moon. Assume that the heat loss from the core is approximately equal to the current production by radioactivity. Compute the expected energy output, using data from Chapter 6 (Table 6.5) as necessary. (b) Then assume that this heat production in the moon's core is just in balance with the heat loss through the ice crust, which has its outer boundary at 94 K and its inner boundary at 170 K (the temperature of the ammonia-rich water sea). Find the thickness of ice that is consistent with this balance.

11.8 Let's consider the possible evolution of a disk around a planet, a structure like one of Saturn's rings. We will suppose that the material is in the form of small particles spread uniformly from an inner radius $R_{\rm i}$ to an outer radius $R_{\rm o}$, with a uniform average surface mass density σ (this is the total amount of matter contained in each square m of the ring plane, and is measured in kg m^{-2}). Within this disk the various particles will be orbiting the central planet (of mass $M_{\rm p}$) in nearly circular Keplerian orbits, but we expect that small deviations from circular orbits will produce collisions between ring particles, causing a kind of friction between neighboring orbits, which will have the effect of gradually slowing the innermost particles (and thus decreasing the inner radius of the disk), and of speeding up the outermost particles (increasing the outer radius). Let's suppose that the disk evolves under the influence of this internal friction but without any external torques acting. (a) Write down an expression for the total mass in a narrow ring of radial extent dR. Integrate this expression from $R_{\rm i}$ to $R_{\rm o}$ to find the total mass $M_{\rm d}$ of the disk in terms of $R_{\rm i}$, $R_{\rm o}$ and σ. (b) Write down an expression for the angular momentum dL of the narrow ring of radial extent dR, and integrate this expression to find the total angular momentum L of the disk. (You may compute the velocity $v(R)$ of the ring at R by equating the gravitational acceleration due to the central planet with the acceleration which produces circular motion.) (c) Write down an expression for the total energy $dE_{\rm tot}$, kinetic plus gravitational (which of course is negative) of a small ring of width dR. Use your expression for $v(R)$ to simplify this result. Integrate the result from $R_{\rm i}$ to $R_{\rm o}$ to find the total energy of the disk (this should be negative). (d) Now suppose that the disk spreads out as a result of the friction discussed above to new inner and outer radii $R'_{\rm i}$ and $R'_{\rm o}$, still keeping the same total mass and uniform surface density. This will decrease the surface density to a new value σ'. Find an expression for σ' in terms of σ and the various R's. (e) Because we have assumed no external torques, the expanded disk should have the same value of L as before. Use this requirement to find one relation between the original values of R_i and R_o and the new values. Don't forget that the expanded disk has a different σ than the original one. [Hint: simplify your result using the fact that $a^4 - b^4 = (a^2 - b^2)(a^2 + b^2)$.] (f) Now find an expression for the new total energy E'_tot of the expanded disk. Use your result from (c), replacing the σ' by your expression from (d). This will result in an expression in both the old and the new dimensions. (g) Now (finally!) suppose that the original disk extended from $1.5R_\mathrm{p}$ to $1.7R_\mathrm{p}$, and that the new inner boundary is at $1.3R_\mathrm{p}$. Use your result from (e) to find the new outer radius of the disk, and then find the *ratio* of the original total disk energy to the new total energy. The result should be a little larger than 1. Since both energies are negative, this shows that the new disk has *decreased* in total energy; this is where the energy comes from that is dissipated in the disk's internal friction. You have just demonstrated a specific example of a general result, that dissipation resulting in expansion decreases the total energy of a disk, and thus is the way in which the disk will evolve. The disk cannot evolve from a wider to a narrower shape without external energy input.

11.9 Suppose a small spherical icy body is held together mainly by solid body crystal forces rather than gravity, and that it has a tensile strength of 1.5×10^6 N m^{-2} (i.e. it is an iceberg and a force per unit area of 1.5×10^6 N m^{-2} would break it). Could any planet in the solar system disrupt it by tidal forces? Assume a density of 1000 kg m^{-3} and a radius of 1 km for the small body.