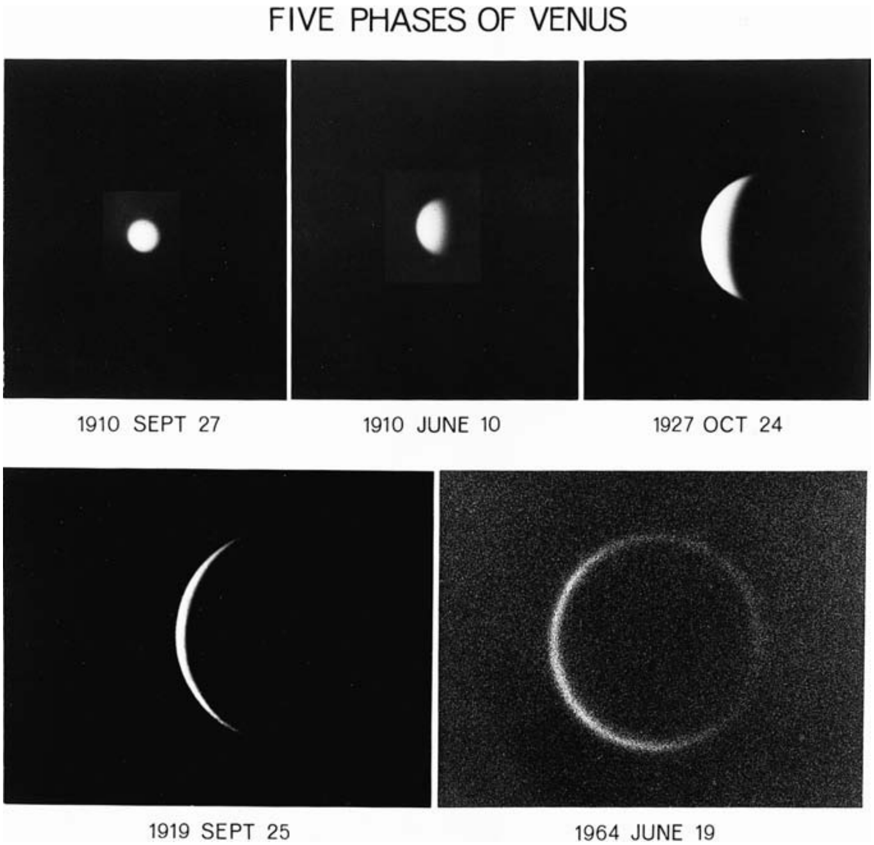


Fig. 7.1 The phases of Venus When fully illuminated, Venus looks small and far away; its apparent size is about seven times larger when the crescents are narrowest and Venus is nearest the Earth. After observing similar phases with his small telescope in 1610, Galileo wrote "Cynthiae figuras aemulatur mater amorum", or "The mother of love (Venus) emulates the figure of Cynthia (Moon)". The phases and variation in the apparent size of Venus provided important early evidence that the planets revolve around the Sun rather than the Earth. (Lowell Observatory photographs.)

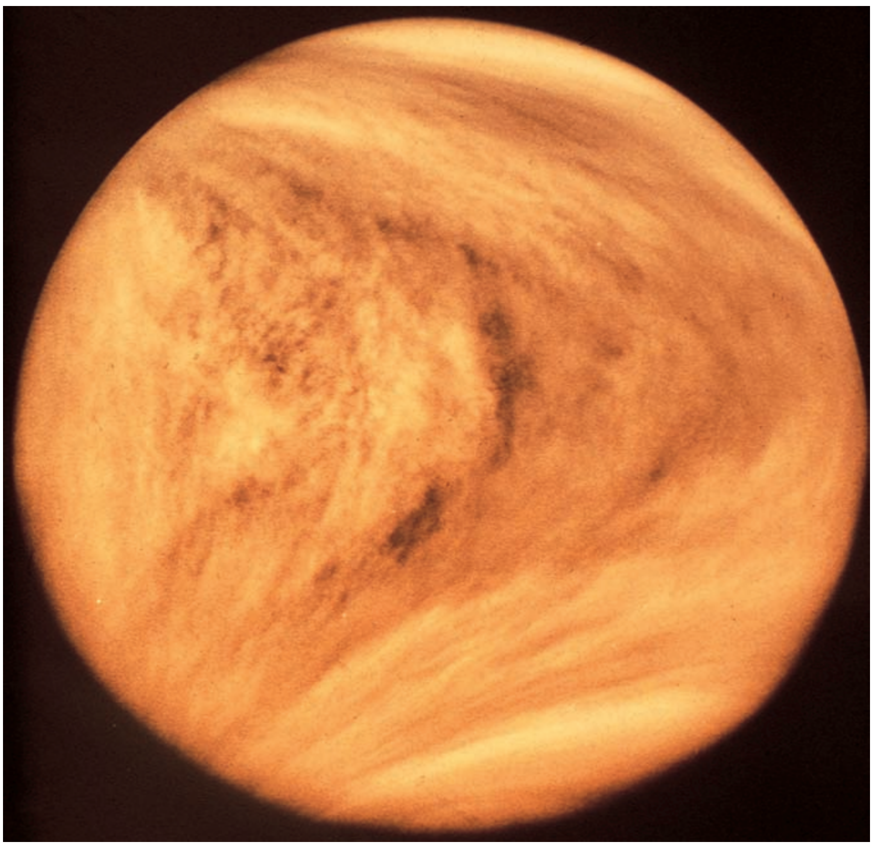


Fig. 7.2 Cloud tops of Venus Clouds of sulfuric acid permanently shroud Venus' surface from view. The horizontal Y-shaped clouds near the equator move from east to west along with high-speed winds at velocities of up to 100 meters per second, circling the planet every four Earth days. This image was taken from the *Mariner 10* spacecraft when it encountered Venus on 5 February 1974, using its gravitational influence to be flung on to Mercury. Similar features have been seen in violet or ultraviolet light from the *Mariner 10*, *Pioneer Venus*, *Galileo*, and *Venus Express* spacecraft, as well as from the *Hubble Space Telescope*. (Courtesy of NASA.)

7.3 Penetrating the clouds of Venus

Venus has a hot and heavy atmosphere

In many ways, Venus is the Earth's twin sister, with almost the same weight and waistline. Its mass is 19 percent less than the Earth's mass, and its mean radius of 6051.8 kilometers is just 5 percent less than the Earth's radius. So the feel of gravity on the planet's surface is like that on Earth. The two planets are also alike in mass density and composition, so at formation they must have been very similar to each other. And because Venus is just a little closer to the Sun than Earth, the climate beneath Venus's clouds was once thought to be warm and temperate.

Since no one could see the surface under the clouds, scientists were free to speculate about what might be found there, and fascinating creatures were imagined to flourish in the warm, wet environment. Throughout the 19th and most of the 20th century, our sister planet was pictured with a verdant, swampy surface, perhaps with oceans of water. It might even be teeming with life.

But space-age scientists have drastically altered this romantic vision. Their radio telescopes and space probes have penetrated the clouds and uncovered the elusive planet's hidden secrets. They revealed a truly hellish and sterile surface, without a trace of flowing water or a sign of life. Beneath her pure, gleaming clouds, the planet of love is a torrid inferno!

The first hint of the inhospitable environment came in 1958 from ground-based radio astronomers when they measured unexpectedly large amounts of microwave radiation emitted by Venus. If the radiation was coming from the planet's surface, it had a temperature of hundreds of kelvin, hotter than a microwave oven turned up high and hot enough to boil away any oceans on Venus.

This explanation was not universally accepted. Some scientists thought the microwave radiation might be coming from the atmosphere, where lightning could generate the emission. But then the United States launched the first interplanetary spacecraft, *Mariner 2*, which flew by Venus in 1962. Instruments aboard *Mariner 2* pinpointed the source of the microwave radiation, showing that it comes from the scorching surface of Venus rather than from its atmosphere. As one scientist put it, Venus is no place to raise the kids.

Diving into the inferno

Of course, the only way to be certain about what lies beneath the clouds was to send a space probe through them. After some initial failures, the former Soviet Union mastered the technique. In 1967 their *Venera 4* spacecraft was the first to enter the atmosphere of another planet, make measurements there, and radio home the results. It confirmed that the atmosphere of Venus is about 96 percent carbon dioxide, and recorded increasing temperatures and pressures on its way down, at least until the spacecraft was either burnt up or crushed to pieces. Venus contains about 300 000 times as much carbon dioxide in its atmosphere as the Earth does.

The day after *Venera 4* made its historic entry into the atmosphere of Venus, *Mariner 5* flew by the planet. Measurements of the way the atmosphere changed *Mariner 5*'s radio signal, when it passed behind the planet and reappeared on the other side, provided a clear profile of the temperature and pressure and confirmed their high values near the ground.

The next two probes, *Veneras 5* and *6*, penetrated further, and then in 1970 *Venera 7* survived the heat and pressure of descent to become the first spacecraft to land on the surface of another planet. *Venera 8* repeated this feat in 1972.

Veneras 7 and *8* measured the atmosphere's temperature and pressure all the way down to the ground, showing that the surface temperature is a sizzling 735 kelvin. That is about as hot as a self-cleaning oven and hot enough to boil the ground dry, and to incinerate any humans that might visit the planet. Venus is the hottest planet in the solar system, with a surface temperature above the average 700 kelvin of Mercury's daytime surface.

The high surface temperature on Venus is maintained by the greenhouse effect, driven by the small percent of solar energy that reaches the ground. The thick atmosphere of Venus traps the Sun's heat near the surface, raising the ground temperature to about three times what it would be without an atmosphere. Although carbon dioxide is the most significant greenhouse gas on Venus, the presence of sulfur dioxide and clouds of sulfuric acid enhance its action.

The massive atmosphere imposes a pressure that is 92 bars, or 92 times the air pressure at sea level on Earth. It would crush you out of existence. The surface pressure is comparable to that experienced by a submarine 500 fathoms, or 1000 meters, below the surface of our terrestrial oceans. So it's a marvel that *Veneras 7* and *8* could withstand the pressure and send back information. None of the landers lasted more than a few hours before succumbing to the destructive heat and pressure.

Five years later, *Veneras 9* and *10* obtained surface photographs, and the Soviets subsequently parachuted seven more entry probes to the surface, determining among other things the composition of the rocks. Two more landers, *Veneras 11* and *12*, descended to the surface in December 1978.

Table 7.2 Some important American, European, and Soviet missions to Venus[a]

Mission	Arrival date[b]	Accomplishments
Mariner 2	14 Dec. 1962	Flyby of Venus, first successful planetary flight, confirmed intense microwaves from surface of Venus, measured solar wind in interplanetary space
Venera 4	18 Oct. 1967	First entry probe of another planet's atmosphere, confirmed that carbon dioxide is its main ingredient
Mariner 5	19 Oct. 1967	Flyby, confirmed high temperature and pressure at surface of Venus
Veneras 5, 6	16, 17 May 1969	Penetrated farther than *Venera 4*
Venera 7	15 Dec. 1970	First probe to land on the surface of another planet, measured surface temperature, pressure and radioactive elements
Venera 8	22 July 1972	Landed on surface
Veneras 9, 10	23, 26 Oct. 1975	First photographs of surface
Pioneer Venus	4 Dec. 1978	First global radar maps of topography, first map of gravity field, five atmospheric probes
Veneras 11, 12	21, 25 Dec. 1978	Landers
Veneras 13, 14	1, 5 Mar. 1982	Chemical analysis of rocks
Veneras 15, 16	10, 14 Oct. 1983	Orbiters, radar images of surface
Vegas 1, 2[c]	11, 15 June 1985	Landers, balloon atmosphere probes
Magellan	10 Aug. 1990	Radar-imaging orbiter, global high-resolution maps of features
Venus Express	11 Apr. 2006	Measurement of atmosphere and cloud patterns, including polar vortex

[a] America's NASA missions to Venus are *Mariner*, *Pioneer* and *Magellan*, Europe's ESA mission is *Venus Express*, and the Soviet missions are named *Venera* and *Vega*.

[b] The *Venera* spacecraft were sent to Venus when the planet was near the Earth, taking about four months to get there.

[c] The main *Vega 1* and *2* spacecraft continued on to comet Halley after dropping probes at Venus.

Two American spacecraft, together known as the *Pioneer Venus* mission, arrived at Venus in the same month as *Venera 12* landed. The *Pioneer Venus Orbiter*, also known as *Pioneer 12*, circled the planet and sent back useful data for 14 years. By sending down radio signals and measuring their echoes, its radar instrument made the first topographic map of the surface, revealing an exceptionally smooth world with just a few high places. Other instruments aboard the *Orbiter* scrutinized the atmosphere, clouds and ionosphere, the electrically charged layer between the atmosphere and outer space.

The *Pioneer Venus Multiprobe*, also called *Pioneer 13*, carried four craft, one large probe and three small ones, which plunged into the atmosphere at both high and low latitudes and on both the daylight and night-time sides, providing a comprehensive picture of the atmospheric structure.

In the late 1960s, the Soviet Union began to hurl spacecraft toward Venus at 19-month intervals, every time the planet's orbital motion carried it nearest to the Earth (Table 7.2). A spacecraft launched near that time requires the least energy and fuel to reach Venus, taking about four months. The Soviets sent them in steadily increasing numbers for two decades; the American launches were fewer, but they included more technologically sophisticated instruments.

Clouds of concentrated sulfuric acid on Venus

What accounts for the unbroken layer of pale yellow clouds that form a thick impenetrable layer at altitudes of between 50 and 80 kilometers above Venus's surface? The billowing white clouds on Earth are formed when water vapor rises about 12 kilometers into the cold atmosphere. But because of the high surface temperature on Venus, there can be no liquid water on its surface and its atmosphere is extremely dry compared to the Earth. It possesses only a hundred-thousandth as much water as the Earth has in its oceans. If all of Venus's water could somehow be condensed onto the surface, it would make a global puddle only a couple of centimeters deep. So there are no water clouds on Venus, and no water rain.

Detailed study of the sunlight reflected from the uppermost clouds indicates that the reflecting cloud particles have a spherical shape, implying that the particles are liquid droplets rather than ice crystals. Water and other plausible liquids were ruled out because they have the wrong reflecting and refracting properties.

Baffled astronomers found the answer in the 1970s. A combination of spectroscopy and polarimetry, or how the cloud droplets polarize light, showed that the clouds of Venus are composed of concentrated sulfuric acid! That is the same sulfuric acid that is commonly used in car batteries and contributes to the eye-stinging quality of smog in some industrial cities, especially near smelters.

But where does the acid come from? Instruments aboard *Pioneer Venus* showed that it is derived from gaseous sulfur dioxide and water vapor in the atmosphere. Chemical reactions involving sulfur dioxide (SO_2) and water vapor (H_2O) form the sulfuric acid (H_2SO_4). A similar series of chemical reactions forms terrestrial smog.

When the concentrated sulfuric acid droplets fall into the warmer atmosphere on Venus, they evaporate, and the resulting gas rises again to the cloud layers. Thus, the acid rain on Venus never reaches the surface, and it is not removed from the atmosphere.

The *Pioneer Venus* investigations also showed that related substances contribute to the greenhouse effect on Venus. Although carbon dioxide is the most significant greenhouse gas on the planet, its action is enhanced by the presence of sulfur dioxide, water vapor, carbon monoxide, and cloud particles. This mixture prevents most of the heat radiation from escaping into space, yielding the torrid surface temperatures.

Venus is certainly too hot for liquid water to exist on its surface, but large amounts of water vapor are bound up with the sulfur dioxide in its clouds. Volcanoes probably supply these gases and maintain the clouds. Solar ultraviolet radiation will decompose any water vapor or sulfur dioxide molecules that rise to high levels in the atmosphere. The hydrogen is continuously lost into space, after water is torn apart. Sulfur and oxygen are too heavy to escape and they react with other atmospheric constituents. The thick clouds that are observed today could only be present if volcanoes supplied their constituents within the past 30 million years. Such volcanoes probably stay active for tens of millions of years, so they may still be active today.

On Earth, water rain efficiently removes sulfur dioxide and other sulfur gases from the atmosphere, so they are only present in very small amounts. Any sulfur gases that are discharged into the air by factories or volcanoes dissolve in our white clouds of water ice, to form droplets of sulfuric acid that are quickly washed out by water rain, sparing us the corrosive acid clouds on Venus. But the sulfuric acid that does rain down to the Earth's surface can damage forests and lakes, so it is still of concern.

Glimpse of a volcanic surface on Venus

Veneras 9 and *10* touched down in October 1975, surviving long enough to send back one picture each, the first from the surface of any other planet. Six years later, *Veneras 13* and *14* parachuted down to the surface and transmitted more images, and analyzed the rocks. Since the sulfuric-acid clouds reflect most of the incident sunlight, and absorb almost all of the rest of it, Soviet astronomers once thought that there might not be any light at the surface. They therefore equipped their earlier spacecraft with floodlights to illuminate the scene when they reached the ground, but it turned out that there was enough natural light to take the historic pictures.

The surface of Venus is always bathed in a diffuse light under a heavy overcast. So the landers found that daylight on Venus resembles a dark, smoggy day in Los Angeles or Mexico City. The pictures are sometimes colored orange, for this is the main color to make it through the thick clouds.

Some scientists thought that the high temperatures and pressures would melt, flatten and chemically weather the surface into a featureless plain. However, the surface photographs showed fresh-appearing rock without eroded edges (Fig. 7.3). The sharp-edged, slab-like rocks are probably formed from flowing lava. As the molten lava spread across the surface and cooled, it would produce the thin, fractured layers of rock that we see in the surface photographs.

Analysis of the surface rocks showed that they have a composition and mass density that resemble basalt, the type of dark, fine-grained lava that lines the Earth's ocean floors and fills the mare basins on the Earth's Moon. The basalt rocks were identified at several landing sites on Venus, suggesting that much of the planet is encrusted by lava, covering its original surface. The basaltic crust was most likely extracted from a differentiated interior, when molten rock rose up through the crust, forming volcanoes that once poured lava over much of the planet's surface.

Circulation of the atmosphere of Venus

At ground level, the heavy atmosphere is sluggish and turgid, hardly moving at all. The speeds measured from the *Venera* landers and the *Pioneer Venus* probes are between 0.3 and 1.0 meters per second, or about the walking speed of a tired old man. A fast wind would have a tremendous

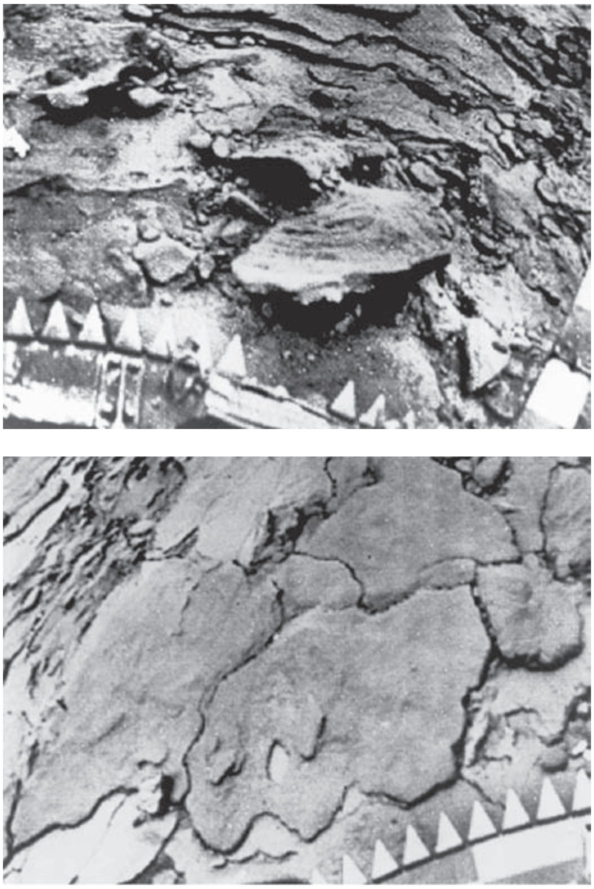


Fig. 7.3 Surface rocks on Venus On 5 March 1982 the *Venera 14* lander touched down on Venus at 13 degrees south latitude and 310 degrees east longitude, where it survived for just one hour before succumbing to the planet's heat. That was long enough to radio back these photographs of the surface of Venus, which include part of the lander and a mechanical arm at the bottom. The thin, plate-like slabs of rock could be due to molten lava that cooled and cracked. The composition and texture of these rocks is similar to terrestrial basaltic lava. (Courtesy of Iosif Shklovskii.)

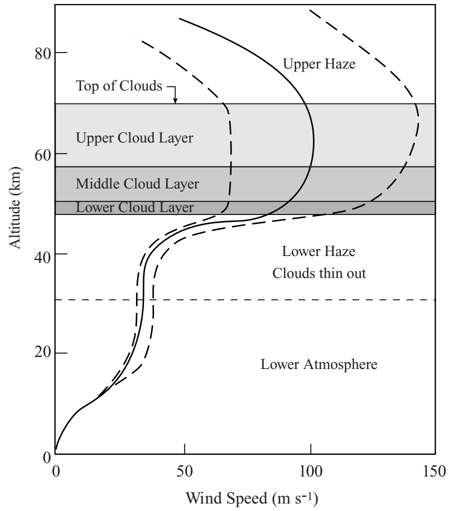


Fig. 7.4 Wind speeds and cloud layers of Venus At all altitudes in the thick atmosphere of Venus, the dominant winds blow westward with a speed that increases with height. From a gentle breeze near the ground, the wind speed increases to 100 meters per second at great heights. Space probes that penetrated the planet's clouds have been able to detect three distinct layers in the opaque sulfurous clouds. The top layer contains small droplets of sulfuric acid; the middle layer contains larger but fewer particles. The bottom layer is the densest and contains the largest particles; it is comparable to bad city smog in visibility. Beneath the lowest layer, the atmosphere is hot enough to vaporize all particles, so it is relatively clear down there.

force in the dense atmosphere, disrupting the surface. Because of the slow winds, the bottom of the atmosphere moves along with the surface.

The wind speed increases with altitude, rising to about 100 meters per second in the clouds at about 60 kilometers in height (Fig. 7.4). The high-flying clouds race around the planet once every four Earth days, from east to west in the same backwards direction that the planet rotates. So the top of the atmosphere is blown around Venus more than 100 times faster than the planet rotates; such a rapid motion is sometimes called super-rotation. These high-speed zonal (east–west) winds are partly driven by the rotation of the solid planet beneath them, but the exact mechanism for maintaining the flow is not well understood.

The atmosphere and winds have transformed the impact craters on Venus, which are unlike those seen on any other world. The dense atmosphere affects the impact debris, changing it into fluid-like flows, and the material ejected during impact is moved by the winds. Some fresh craters are surrounded by radar-bright halos, streamlined hoods and tail-like wind streaks that act like wind vanes, pointing downwind at the time of impact (Fig. 7.5). The wind streaks indicate that the winds just above the surface were blowing toward the equator from the northern and southern hemisphere.

The atmosphere redistributes heat from one part of Venus to another, thereby moderating temperature differences. Most of the sunlight falling on Venus is either reflected by the clouds or absorbed in them. And because the Sun's rays fall directly on the equator and obliquely at the poles, the equatorial clouds are initially warmer than the polar ones. But this temperature difference generates winds that transfer heat in a single large Hadley cell (Fig. 7.6), named after George Hadley (1685–1768) who first proposed such a circulation for the Earth's atmosphere.

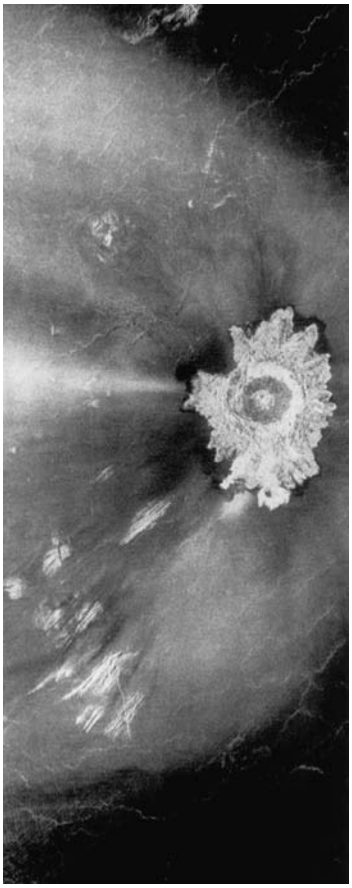


Fig. 7.5 Winds blow material ejected from a crater on Venus A radar image of Avidar Crater is surrounded by streamlined, horseshoe or parabolic shapes (*top and bottom right*) and a bright, jet-like streak (*center left*) that extends over the surrounding plains. These unusual features, seen only on Venus, are believed to result from the interaction of ejected debris and high-speed winds in the upper atmosphere, blowing from the east (*right*). The crater is 30 kilometers in diameter, and is named for the Turkish educator and author Halide Adivar (1883–1964); it is located at 9 degrees north latitude and 76 degrees east longitude, just north of the western Aphrodite highland. (Courtesy of NASA/JPL.)

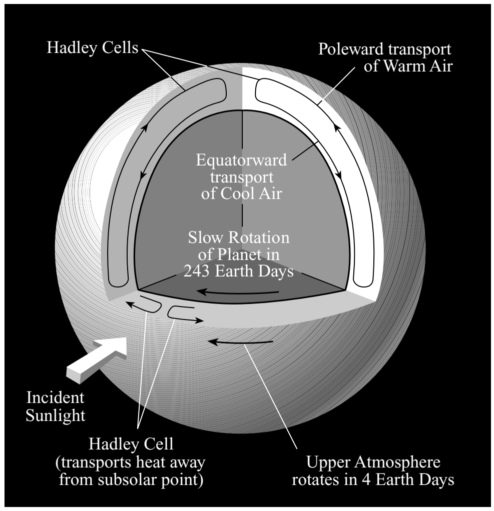


Fig. 7.6 Hadley cells in the atmosphere of Venus Incident solar energy drives a Hadley-cell circulation of the atmosphere, which keeps heat from building up at any one location. Air rises in warm regions near the equator on the sunlit side of Venus, where the planet is heated most by the Sun. Some of the warm air flows towards cooler zones near the poles, and sinks and returns to warm equatorial regions again. Strong winds also blow around the planet in the direction that Venus rotates, from east to west.

As on the Earth, the warm air rises at the equator to the cloud tops, where winds propel it toward both poles. After warming the poles, the circulating atmosphere sinks and flows back towards the equator at lower levels near the base of the clouds, completing the cell. The stronger zonal (east to west) circulation on Venus combines with this weaker Hadley (north–south) circulation, giving rise to a wind vortex that carries the clouds in a slow spiral toward the poles.

The turbulent atmosphere of Venus has been studied in great detail using ultraviolet and infrared detectors aboard the *Venus Express* spacecraft, which began orbiting the planet on 11 April 2006. These instruments have provided new insights to the planet's global cloud patterns that were first detected at ultraviolet wavelengths. The equatorial areas on Venus that appear dark in ultraviolet light are regions of relatively high temperature, where intense convection brings up dark material from below. In contrast, the bright regions at mid-latitudes are areas where the temperature in the atmosphere decreases with depth. The *Venus Express* observations have also confirmed that lightning is produced in the sulfuric acid clouds.

At the polar regions, gases that were moving toward the poles are deflected sideways by the planet's rotation, creating a huge whirling vortex that resembles the eye of a gigantic hurricane (Fig. 7.7). An immense, spinning vortex spirals around both poles. Each vortex draws the atmosphere downward, like water running down a bathtub drain, and they change shape in a matter of days, even forming a double eye in one polar region on occasion.

The dense lower atmosphere transports heat so efficiently from one part of the globe to another that the ground temperature varies by no more than a few degrees between the equator and poles or from day to night, so there is no place to escape the heat.

The changing distribution of sunlight on the Earth causes stormy weather and produces the changing winds that propel clouds across the globe. In contrast, Venus has a steady atmospheric circulation pattern almost devoid of

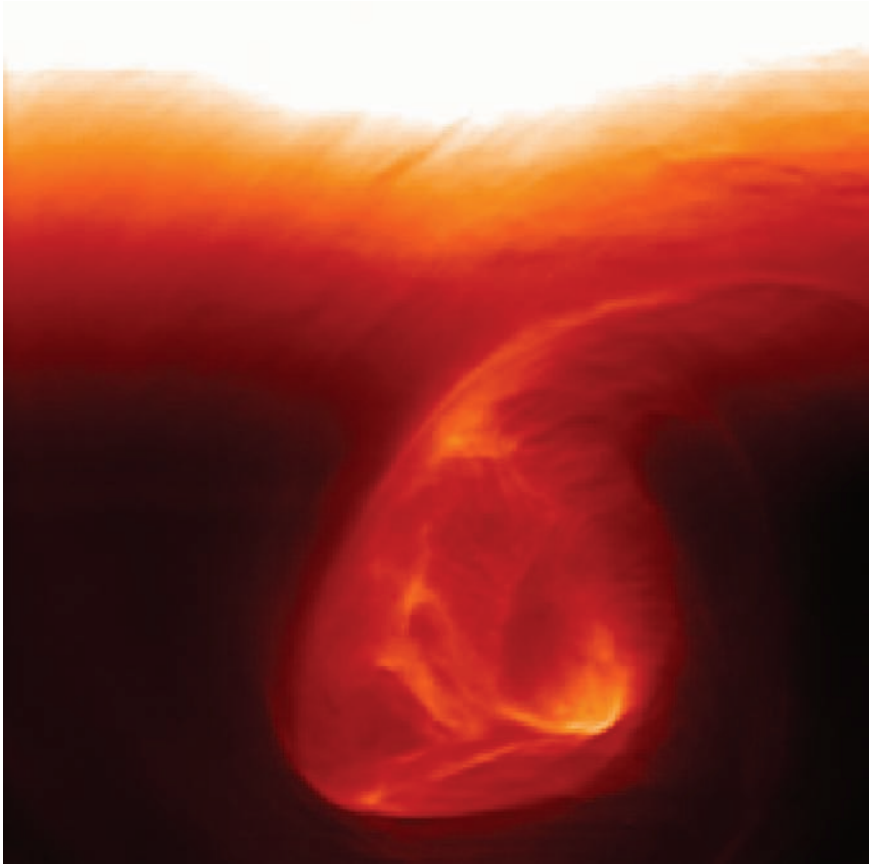


Fig. 7.7 South polar vortex at Venus
The core of a whirling polar vortex shines brightly in this infrared image taken on 9 August 2007 from the *Venus Express* spacecraft. The vortex, which resembles the eye of a hurricane on Earth, is located near the south pole of Venus. The 2000-kilometer-wide vortex was first discovered from the *Mariner 10* flyby of Venus on 5 February 1974. There is a similar structure near the planet's north pole, which was observed from *Venus Express*. The central part of the southern vortex is very dynamic, changing shape within a matter of days, including the formation of a double-eye. (Courtesy of ESA.)

weather. Since the planet's orbit is nearly a perfect circle and its rotation axis is hardly tipped at all, there are no noticeable seasons on Venus.

Missing magnetism on Venus

Unlike the Earth, Mercury and the four giant planets, Venus has no significant global magnetic field, for reasons not fully understood. Measurements from *Pioneer Venus Orbiter* showed that if Venus has any magnetic field its strength is less than one hundred-thousandth that of Earth, or less than 3×10^{-9} tesla. Lacking a magnetic field, Venus has no belts of trapped particles such as the Van Allen belts near Earth.

The weakness of this magnetic field is rather surprising because Venus and the Earth have a similar size and mass, and they might be expected to have similar interiors. The Earth has a molten outer core and a solid inner one, and by analogy Venus ought to possess similar cores. However, Venus does not show the magnetic field that would be produced by currents within a molten outer core.

Contrary to popular belief, the slow rotation of Venus is not responsible for the lack of a significant magnetic field, but something in its core is out of whack. Some investigators argue that the core is now completely solid, while others suggest that core solidification has not yet commenced. In either case, Venus would lack both a molten outer core for currents to flow in and the heat given off by the creation of a solid inner core, which would help maintain the flow.

The planet's lack of an appreciable magnetic field exposes the upper atmosphere to the continuous hail of charged particles from the Sun. The intrinsic magnetic fields of the Earth and other planets fend off and deflect this electrically charged solar wind. Nonetheless, the solar wind is prevented from reaching the surface by Venus's dense atmosphere and ionosphere.

Energetic ultraviolet sunlight ionizes some of the atoms and molecules in the outer atmosphere above the clouds, forming an electrified layer similar to the Earth's ionosphere, and this layer helps shield the ground from the solar wind. The ions provide conduction paths for electrical currents that produce forces that counter the wind. As a result, the solar wind slows down and is deflected around the planet in a bow shock, and the interplanetary magnetic field is draped back to form a magnetotail (Fig. 7.8). The solar wind also accelerates the charged, ionized particles in the upper atmosphere of Venus, giving some of them enough energy to escape the planet. So Venus slowly loses

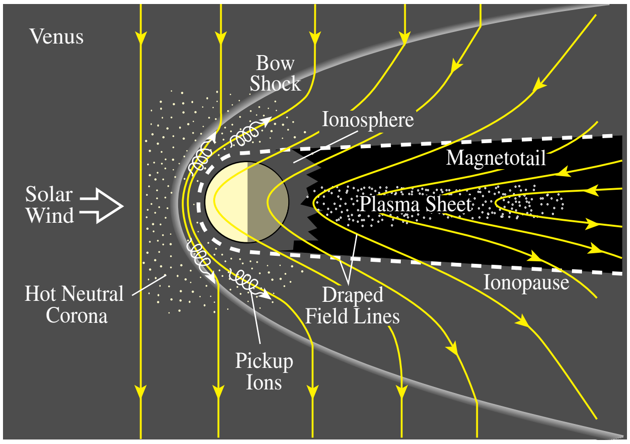


Fig. 7.8 Solar wind flows around Venus Even though Venus has no appreciable magnetic field, the solar wind is prevented from reaching the surface by Venus' dense atmosphere and by electrical currents induced in its conducting ionosphere. The planet has a well-developed bow shock, but it does not have belts of trapped particles.

a small fraction of its atmosphere due to interaction with the solar wind (Fig. 7.9).

7.4 Unveiling Venus with radar

Venus rotates backwards, at an exceptionally slow rate

Although no human eye has ever seen the surface of Venus, radio waves can penetrate its obscuring veil of clouds and touch the landscape hidden beneath. By bouncing pulses of radio radiation off the surface, radar astronomers discovered in 1967 that Venus spins in the backward direction, opposite to that of its orbital motion. That is, unlike the other terrestrial planets, Venus does not rotate in the direction in which it orbits the Sun.

The radar observations also showed that Venus spins with a period longer than any other planet, at 243.018 Earth days. This rotation period is even longer than the planet's 224.7 Earth-day period of revolution around the Sun, so the

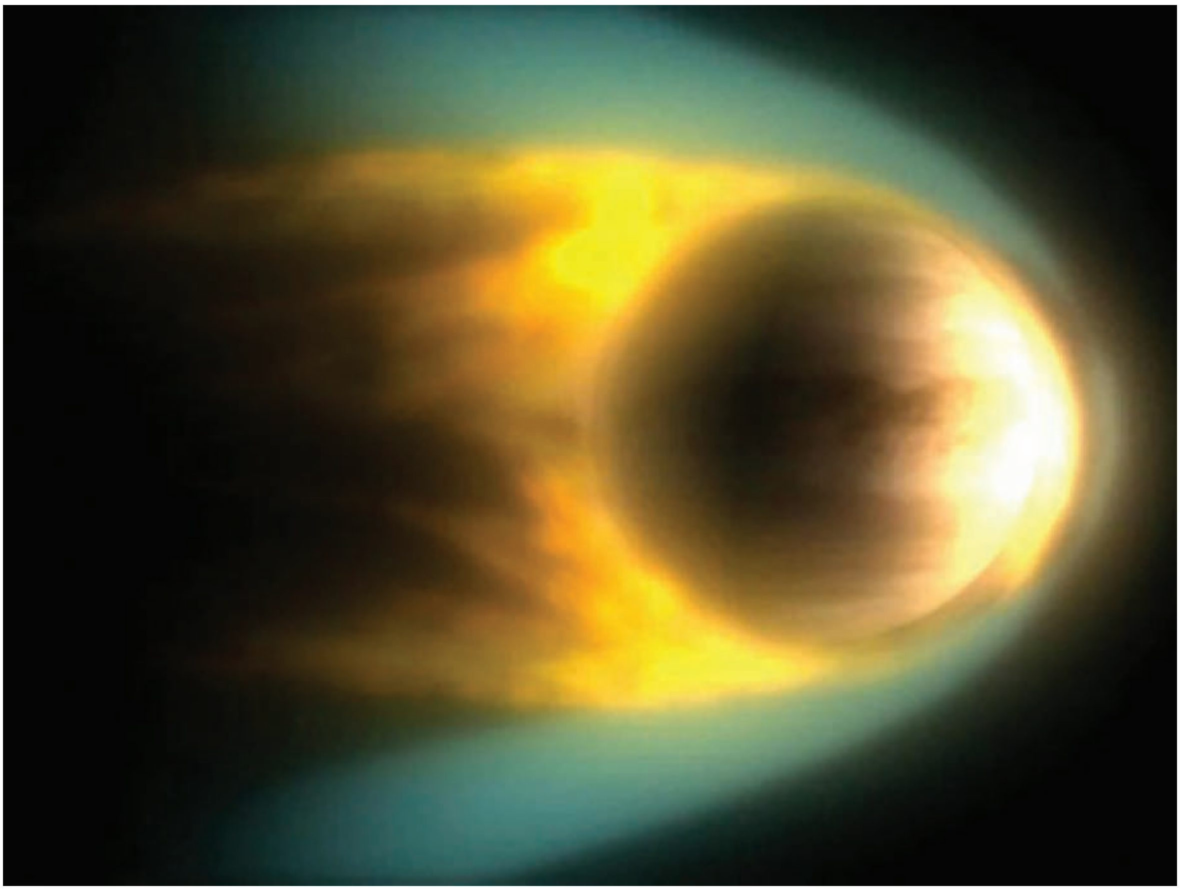


Fig. 7.9 Venus slowly loses some atmosphere Venus is almost as large as Earth, and it is difficult for its atmosphere to escape from the planet's gravitational pull. However, the upper part of the atmosphere on Venus is ionized by solar radiation, and the solar wind accelerates some of these ions, giving them enough energy to escape and join the wind. In this way, the solar wind interacts with the thick atmosphere, and in effect pulls some of it into space. (Courtesy of ESA.)

day on Venus is longer than its year. The method of determining this slow rotation is the same Doppler technique used to discover Mercury's slow 58.6 Earth-day rotation period, and both planets have probably been slowed down by the Sun.

Tides raised by the Sun in the planet's thick atmosphere may explain why Venus turns very slowly and in the wrong way. The Sun's gravitational force produces two tidal bulges; the rotating planet drags these bulges along with it, causing them to twist out of alignment with the Sun. As a result, the Sun's gravitational attraction tends to oppose the rotation, slowing the planet down. Friction between the atmospheric tides and the planet's surface also helped to gradually slow Venus's rotation to the point where it stopped and reversed. The friction of the tides that our Moon produces in the Earth's oceans similarly slows down our planet's rotation.

Magellan and its predecessors reveal a volcanic surface on Venus

The surface of Venus cannot be seen at the visible wavelengths used by our eyes; it can only be sensed by radio transmissions. Radar, an acronym for radio detection and ranging, uses its own source of radio radiation, and does not need sunlight to probe the planet, gathering data day or night. Only radar is capable of piercing the thick clouds of sulfuric acid that blanket Venus.

The planet's topography is inferred from the length of time it takes for a radar pulse to reach a particular part of the surface and return an echo. The surface roughness is determined from the strength of the echo. Rough surfaces and slopes tilting toward the radar reflect more energy and return a stronger echo, thus appearing bright in a radar image. Surfaces that are smooth or tilt away from the incoming radio signal send less energy back, and appear dark in a radar image, somewhat like a wet road seen in the headlights of a car at night.

The surface features on Venus have been probed with radar systems of increasingly greater resolution over the decades. Blurred images were first obtained from powerful, ground-based radio telescopes, such as the one in Arecibo, Puerto Rico. The *Pioneer Venus Orbiter* then zoomed in to take a closer look in the 1970s, acquiring a global radar map of Venus at a coarse resolution of about 100 kilometers. This was followed by the twin spacecraft *Veneras 15* and *16*, whose radar instruments mapped much of the planet's northern hemisphere in the mid-1980s, resolving features as small as 1 kilometer across.

The radar system aboard the orbiting *Magellan* spacecraft next scanned Venus for more than four years in the early 1990s. It mapped details as small as 120 meters across, producing the most complete global view available for any planet, including Earth (Fig. 7.10). The spacecraft is named after the Portuguese explorer Ferdinand Magellan (1480–1521) whose expedition first circumnavigated Earth.

Magellan's radar images revealed a rich and varied landscape with stunning and unprecedented clarity, describing a surface whose nature and history turned out to be quite different from those of the Earth. The surface of Venus is shaped largely by volcanic activity. Its topography is dominated by massive, global outpourings of lava, punctuated by numerous shield volcanoes and unique volcanic constructs never seen before, scarred by sparse, pristine impact craters surrounded by beautiful outflows, and fractured, stretched, crumpled and split open by upwelling magma. Venus has more volcanoes than any other terrestrial planet. With the exception of Jupiter's moon Io, Venus is the most volcanic world in the solar system.

Perhaps the most stirring aspect of the *Magellan* images is the fresh, pristine nature of the features they reveal. Although no one expected to see significant evidence of erosion and weathering on the dry planet, observers were struck by the detailed sharpness of its craters, fractured plains, volcanoes and crumpled landmasses. Most of these surface features have been preserved for hundreds of millions of years. Everything is totally exposed and largely preserved — at least between periods of intense volcanic activity or internal upheaval.

As Aladdin said from his flying carpet in the Walt Disney movie *Aladdin*, it is:

> A whole new world.
> A dazzling place, I never knew.
> But when I am way up here,
> It's crystal clear.

Magellan had one instrument, the Synthetic Aperture Radar (SAR), capable of mapping its topography as an altimeter, and measuring electromagnetic radiation from the surface. Most typical radar systems send out one pulsed radio signal at a time, and process each echo by itself before sending out another pulse. In contrast, *Magellan* sent out several thousand radar pulses each second, and its SAR used computers to accumulate multiple echoes received from many locations simultaneously. The combined data simulated a large antenna and provided the superb resolution and fine detail.

As the spacecraft looped around the poles of Venus, the slowly rotating planet turned beneath it, exposing the entire globe to scrutiny during each 243-day rotation. Every second, the computers took in 36 million bits of data and

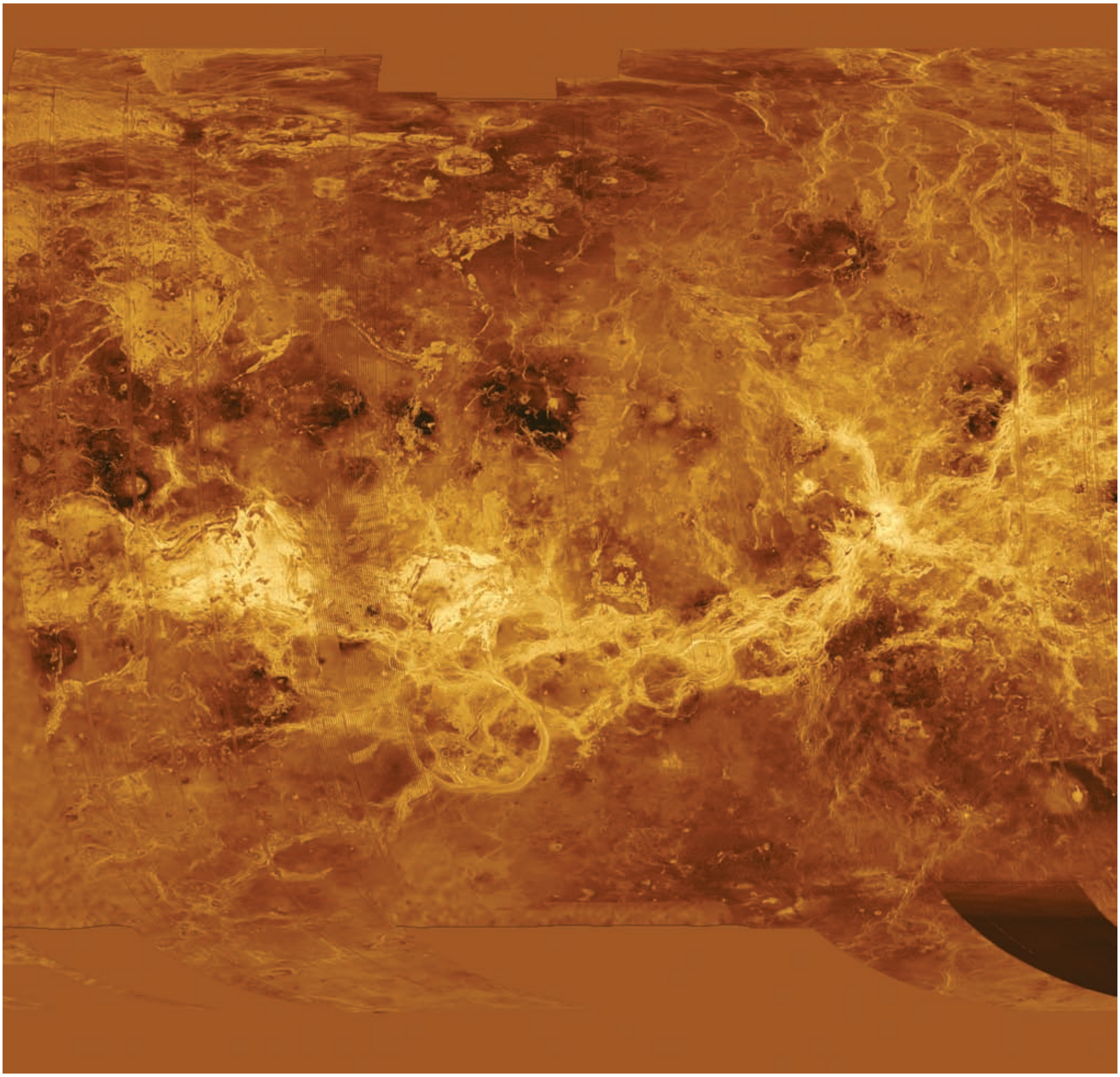


Fig. 7.10 Surface terrain of Venus The cloud-penetrating radar system aboard the *Magellan* spacecraft lifted the perpetual veil of thick clouds on Venus, revealing its surface features in high resolution of 120 meters in the early 1990s. The eastern half of the planet's surface is portrayed in this cylindrical map, extending between 52 and 240 degrees east longitude (*left side to right side*) and from 90 degrees north latitude to 90 degrees south latitude (*top to bottom*). (Courtesy of NASA/JPL.)

relayed them back to Earth. Over four years, they obtained more than a million billion bits of information, exceeding the amount that had been obtained from all previous lunar and planetary spacecraft combined.

To complement the wide-angle, side-looking antenna, there was a second, narrow-beam antenna that looked straight down and used a pulsed signal and its echo to measure the topography. Elevations were measured with an accuracy of 30 meters with a resolution of about 5 kilometers.

Venus is a smoothed-out world

Radar data from the *Pioneer Venus Orbiter* and *Magellan* spacecraft showed that Venus is an extraordinarily smooth world, largely at one low level and quite different from the Earth (Fig. 7.11). Without its water, the topography of the Earth occurs at two distinct elevations, which correspond to the continents and ocean floors. In contrast, about 85 percent of the surface of Venus lies within one kilometer of the average planetary radius, 6051.8 kilometers.

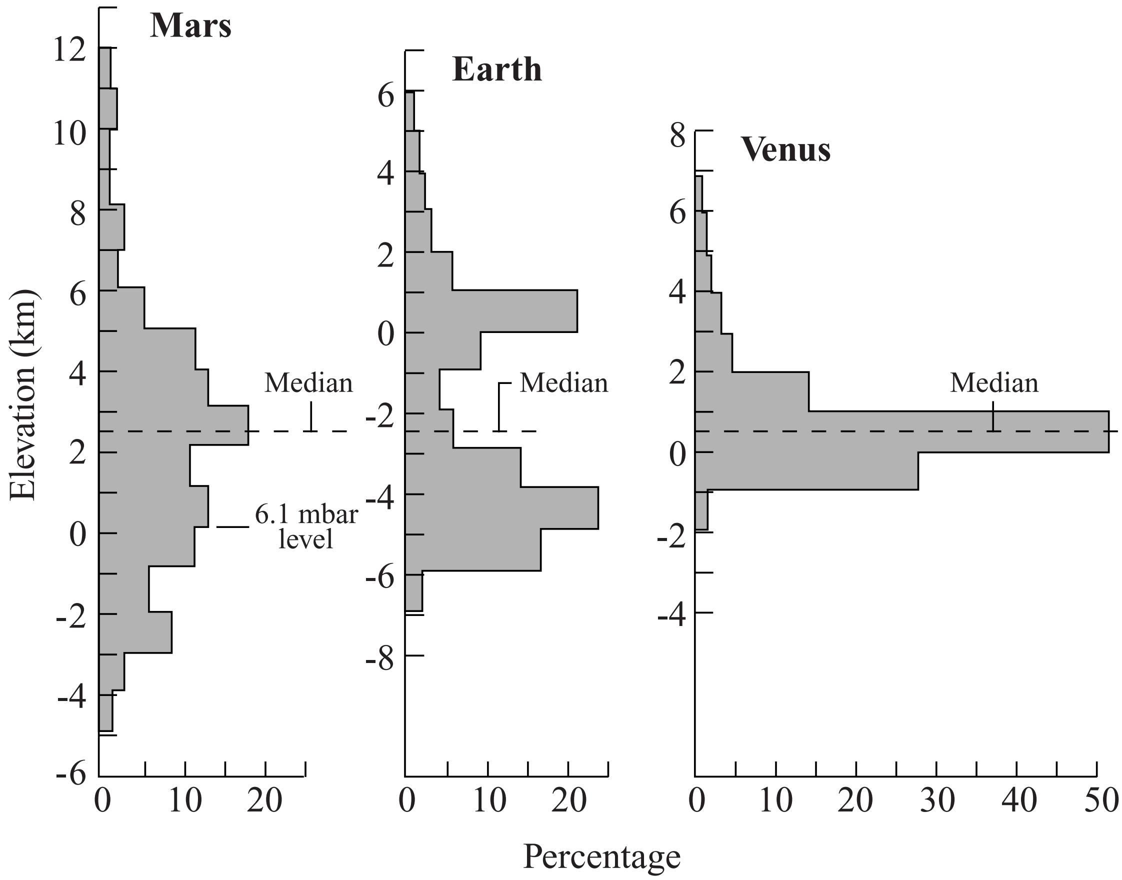


Fig. 7.11 Surface elevations for Earth, Venus and Mars Most places on Earth (*center*) stand near one of two prevailing levels, the high-standing continents or the low-lying ocean floors. In contrast, the surface of Venus (*right*) is unusually smooth and flat; about 60 percent of its terrain lies within 500 meters of the mean planetary radius of 6051.8 kilometers. A small percentage of its terrain consists of elevated highlands that are comparable in height to many continents on Earth. The surface features on Mars (*left*) are spread over a broader range of elevation than most of those on Venus; but the Martian surface elevations are not double-peaked as on Earth.

A coating of lava has smoothed these vast low-lying plains.

The lowest point of Venus is about 6048.0 kilometers and the highest point is at 6062.57 kilometers. Thus the variation of topography on Venus is almost 14.6 kilometers. The surface temperature and pressure varies with height, between 653 and 766 kelvin and between 45 and 119 bars at the highest and lowest elevations.

Although most of Venus's terrain consists of smooth, low-lying, volcanic plains, about 15 percent of the planet's surface consists of highlands that tower above the plains, rising an average 4 to 5 kilometers above the mean planetary radius. There are two large-scale elevated regions that punctuate the smoothed-out surface; they are Aphrodite Terra in the equatorial region and Ishtar Terra in the far north (Fig. 7.12).

Aphrodite Terra is over 10 000 kilometers long and covers a quarter of the planet's circumference at the equator. It contains tall volcanoes, long lava flows, and deep faults and fractures. Western Aphrodite is built from the massifs Ovda and Thetis Regiones; Atla Regio occupies the eastern part of Aphrodite. Ishtar Terra fills about half the planet's circumference at its high northern latitudes and is about the size of Australia. It consists of an elevated plateau encircled by narrow mountain belts (Fig. 7.12).

Each feature on Venus is described by two names — a woman's name that identifies it and a feature designation (Focus 7.1). As examples, Aphrodite and Ishtar are respectively the Greek and Babylonian goddesses of love, and a terra is an extensive landmass. The name of a goddess or a famous mortal woman identifies almost every feature on Venus. The only exceptions on Venus are Maxwell Montes, named in honor of the British physicist James Clerk Maxwell (1831–1879), whose 19th-century theories of electromagnetism made radar possible today, and Alpha Regio and Beta Regio, using the first two letters

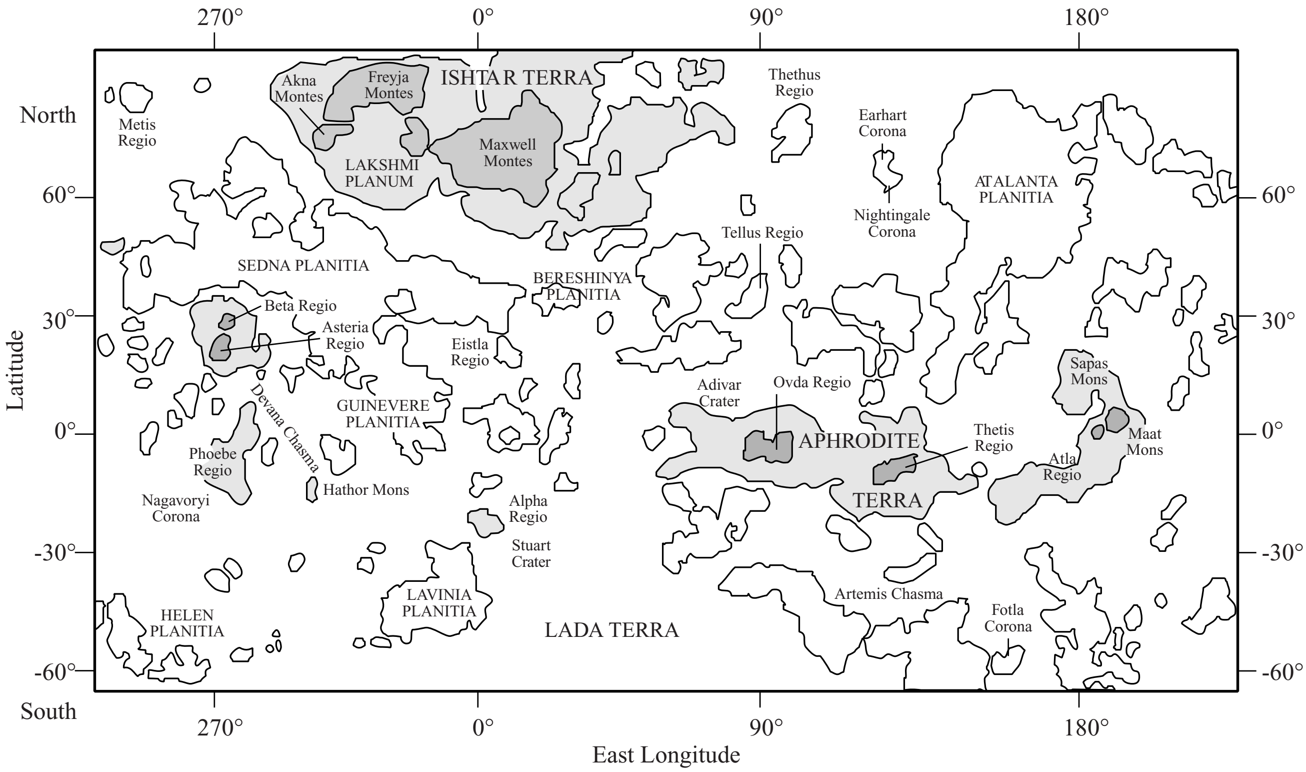


Fig. 7.12 Major surface features on Venus Most of Venus lies at roughly the same radius, in the lowland, volcanic plains (*white areas*). The highland massifs (*dark areas*) include Aphrodite Terra, an elongated feature extending along the equator between 70 degrees and 210 degrees east longitude. Many of the elevated regions near the equator mark the sites of volcanism, such as Maat Mons and Sapas Mons (see Fig. 7.16). The other main elevated region is Ishtar Terra in the north at about 0 degrees longitude (*center top*). Ishtar Terra is roughly the size of Australia. The elevated plateau in the western part of Ishtar Terra is known as Lakshmi Planum, which is bounded on three sides by mountains (see Fig. 7.17), including Maxwell Montes. This mountain's 11-kilometer height above the average radius exceeds by two kilometers the height of Mount Everest above sea level.

Focus 7.1 Naming features on Venus

There are two names that describe every feature on Venus. They are a particular name that identifies it, plus a descriptive name that says what it looks like. As indicated in Table 7.3, each class of feature has a specific type of associated identifying female, either real or mythological.

The identifying names can be suggested by anyone, but eventually they have to be approved by the International Astronomical Union (IAU). The lists of approved names for the planets and their satellites are fascinating, and you can review them in the IAU's Gazetteer of Planetary Nomenclature located on the web at http://planetarynames.wr.usgs.gov/.

of the Greek alphabet. They were first seen in Arecibo radar images in the 1960s, before the female naming convention.

Craters larger than 20 kilometers in diameter are identified by the last name of a famous woman, who has to have been dead at least three years at the time of naming. Some of my favorites include the singer Billie Holiday, the anthropologist Margaret Mead, the writer Flannery O'Conner, the ballet dancer Anna Pavlova, the singer Edith Piaf, and the abolitionist Harriet Tubman. Smaller craters are given common female first names, as are hurricanes on Earth.

Some representative examples for the names of other topographic features include: Fotla Corona, named for a Celtic fertility goddess; Freyja Montes, with the name of a Nordic goddess who weeps gold tears for her lost husband; Maat Mons, with the namesake of an Egyptian goddess of truth and justice; and Ovda Regio, named after a female Finnish spirit who wanders the forest naked to find people to tickle to death.

Table 7.3 Common features on the surface of Venus and the category of women used to identify them

Descriptor term (singular, plural)	Feature type	Category of identifying female
Chasma, Chasmate	Canyon, steep-walled trough	Goddesses of hunt or the Moon
Corona, Coronae	Oval-shaped feature	Fertility and earth goddesses
Crater, Craters	Circular depression	Famous women
Dorsum, Dorsa	Ridge	Goddesses of the sky
Linea, Lineae	Elongate marking	Goddesses of war
Mons, Montes	Mountain	Miscellaneous goddesses[a]
Patera, Paterae	Irregular crater, shallow	Famous women
Planitia, Planitiae	Low plain, level surface	Mythological heroines
Planum, Plana	Plateau or high plain	Goddesses of prosperity
Regio, Regiones	Large area, broad region	Giantesses and titanesses[b]
Rupes, Rupes	Cliffs or scarps	Goddesses of hearth and home
Terra, Terrae	Highland, extensive landmass	Goddesses of love
Tessera, Tessarae	Tile-like, polygonal terrain	Goddesses of fate or fortune

[a] The one exception is Maxwell named after James Clerk Maxwell (1831–1879).
[b] The two exceptions are Alpha Regio and Beta Regio.

7.5 Volcanic plains on Venus

Planet-wide covering of lava on Venus

Venus has been extraordinarily volcanic. Extensive lava flows emerge from cracks in the crust and from towering volcanoes, and rivers of lava snake their way across the planet. Outpouring lava has covered much of the surface of the planet, and tens of thousands of small volcanoes are now found all across its face. Venus exhibits every type of volcanic edifice known on Earth, and some that have never been seen before.

The spreading lava has flooded and filled the low-lying regions of Venus, creating extensive smooth plains that cover about 85 percent of Venus's surface. The volcanic nature of these lowland plains, each designated by the term *planitia*, is demonstrated in the *Magellan* images. You can practically see the molten rock spreading like heavy cream across these plains, often running for hundreds of kilometers down gentle slopes. The magma has risen from within canyons as the crust split open, cooling and solidifying into lava flows that look like frozen river currents (Fig. 7.13).

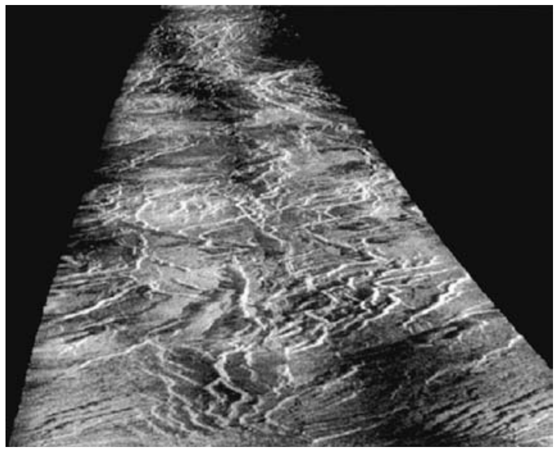


Fig. 7.13 Lava flows This *Magellan* image shows a lava-filled canyon produced when Venus' crust was pulled apart and magma rose within the gap. The lowland plains have many similar canyon systems; typically about 10 kilometers wide and up to 1000 kilometers long, and apparently containing solidified lava flow. This region, located near the equator between Asteria Regio and Phoebe Regio, is about 10 kilometers wide and stretches 600 kilometers to the north (*top*). (Courtesy of NASA/JPL.)

In other places, the molten material has burned paths in the pre-existing lava deposits, following a narrow, sinuous smoothly curving course. They can meander for thousands of kilometers across the planet's surface. Many end in outflows that look like river deltas. These river-like channels were formed not by water, but by lava that was hot enough to carve through solid rock, remaining hot and liquid over distances that are longer than the Nile, the longest river on Earth. The high surface temperature on Venus probably kept the lava liquid, and prevented the cooling flow front from damming up the molten rock behind it.

Fig. 7.14 Stuart Crater This beautiful crater exhibits asymmetric radar-bright ejected material attributed to an oblique impact and interaction with the dense, thick atmosphere of Venus. The crater's rim, which is 67 kilometers in diameter, encircles a bright floor that may have unique physical properties. This impact crater's name honors Mary Stuart, Queen of Scots (1542–1587). (Courtesy of NASA/JPL.)

A relatively young surface on Venus

Venus, like all planets, has been subjected to a continual rain of meteorite bombardment over the eons. The plains of Venus are uniformly peppered with impact craters (Fig. 7.14), the scars of this bombardment, though nowhere near as liberally as on the surfaces of the Moon, Mars, or Mercury. A comparison of craters with the same size on Venus and the Earth's Moon indicates that those on Venus are far fewer in number and more widely spaced than those on the Moon. At one time Venus was probably as heavily pockmarked with large craters as the Moon's ancient surface is, but the scarcity of the craters now on Venus indicates that the surface we now see is much younger than the lunar surface.

We can estimate when the lava flowed by counting the number of craters of a given size on the plains and comparing it to the number on the Moon – ignoring the smallest ones that are not found on Venus because the incoming projectiles burned up in the thick atmosphere. The Moon's crater record tells us the number of impact craters that should appear in a given time, and observations of the craters left on Venus tell us how many have been removed by burial under volcanic flows.

The *Magellan* spacecraft has logged about 1000 impact craters, which when compared to the lunar record indicates an average surface age of about 750 million years, with a large uncertainty of a few hundred million years. Estimates between 300 million and 1 billion years are possible for the global lava covering. At that time molten rock emerged from inside the planet, spreading across the surface, eradicating all previous craters, and wiping out all traces of the first 90 percent of Venus's history. That is a relatively young age, only about 10 percent of the age of the solar system and of the planet Venus itself. No matter how you look at it, the surface of Venus is practically newly born compared with the surfaces of Mars, Mercury and the Earth's Moon, which still bear the scars of a heavy bombardment about 3.9 billion years ago.

Theories for the volcanic makeover of Venus

Everyone agrees that the smooth plains covering most of Venus came from volcanic floods, emanating from the planet's interior, but the experts disagree over when and how it occurred. According to one "global catastrophe hypothesis", planet-wide volcanism wiped the face of Venus clean about 750 million years ago, resurfacing the entire globe and drowning any existing craters in a flood of lava.

There are two equally likely catastrophe interpretations that cannot be distinguished. One is a single resurfacing at about 750 million years ago. The second is that there was continuous resurfacing of the planet over most of its earlier history, and that this resurfacing slowed down sharply at about 750 million years ago. In either interpretation, the planet switched over to a low rate of localized volcanism about 750 million years ago, but it has not disappeared.

Most of the craters on Venus have lava on their floors, and some of them show exterior volcanism affecting their ejecta or breaking their rims, so there has been at least a modest level of ongoing volcanism.

A different model that has observational support is a series of discrete volcanic plains forming events extending over a considerable period of geologic time, with the volume of lava declining from one event to the next. In

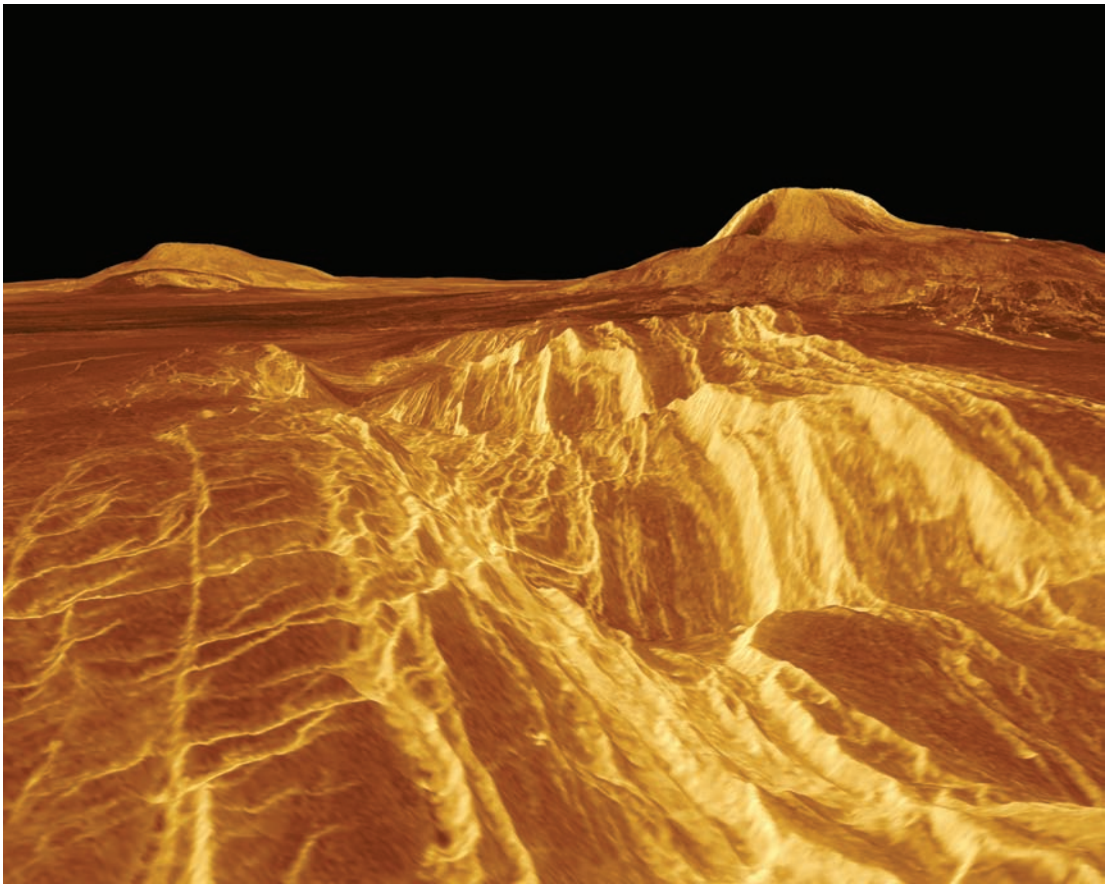


Fig. 7.15 Sif Mons and Gula Mons Two volcanoes are located at the horizon of this three-dimensional *Magellan* radar view of the western Eistia Regio in the central part of Venus' surface, near the equator and zero degrees east longitude. In this computer-generated image, the vertical scale has been exaggerated to show the edge of a rift valley in the foreground. It extends all the way to Gula Mons, the volcano on the right horizon, which stands 3 kilometers high and is located 700 kilometers away. The volcano on the left horizon, called Sif Mons, has a height of about 2 kilometers and a diameter of about 300 kilometers. Eistia is a Norse giantess, Gula is the name of a Babylonian earth-mother and creative force, and Sif is the name of a Teutonic goddess and Thor's wife. The artificial tints are based on color images from the *Venera 13* and *14* landers, simulating the color of sunlight that filters through the thick atmosphere. (Courtesy of NASA/JPL.)

this view, volcanism is an ongoing process, occurring at different places and times, further in the past and in the future from 750 million years ago. In fact, volcanic rises, coronae and rifting suggest that volcanism on Venus may extend up to the present.

The intense volcanism that repaved Venus may have transformed its climate and roasted its air. When massive volcanoes erupted roughly 750 million years ago, they should have ejected a lot of greenhouse gases into the atmosphere in a relatively short time. Calculations indicate that the volcanic release of carbon dioxide, water vapor and sulfur dioxide might have raised the temperatures well in excess of the "present inferno". Still, this volcanism is not the cause of the present intense greenhouse environment on Venus, which probably existed long before the volcanic resurfacing.

7.6 Highland massifs on Venus

Towering volcanoes and mountain ranges

Highland regions in the north and near the equator rise above the low terrain on Venus and account for about 15 percent of the planet's surface. Large-scale plumes of rising magma have probably pushed these elevated regions up from below. When the molten rock pierced the surface, lava flowed out to form towering volcanoes that are perched atop the raised highlands. Some of the volcanoes are found in Beta Regio on the western side of the equatorial highlands. Others are found in the western part of Eistia Regio (Fig. 7.15), located near the center of the equatorial highlands at about zero degrees longitude. Several rise out of Atla Regio in the eastern end of

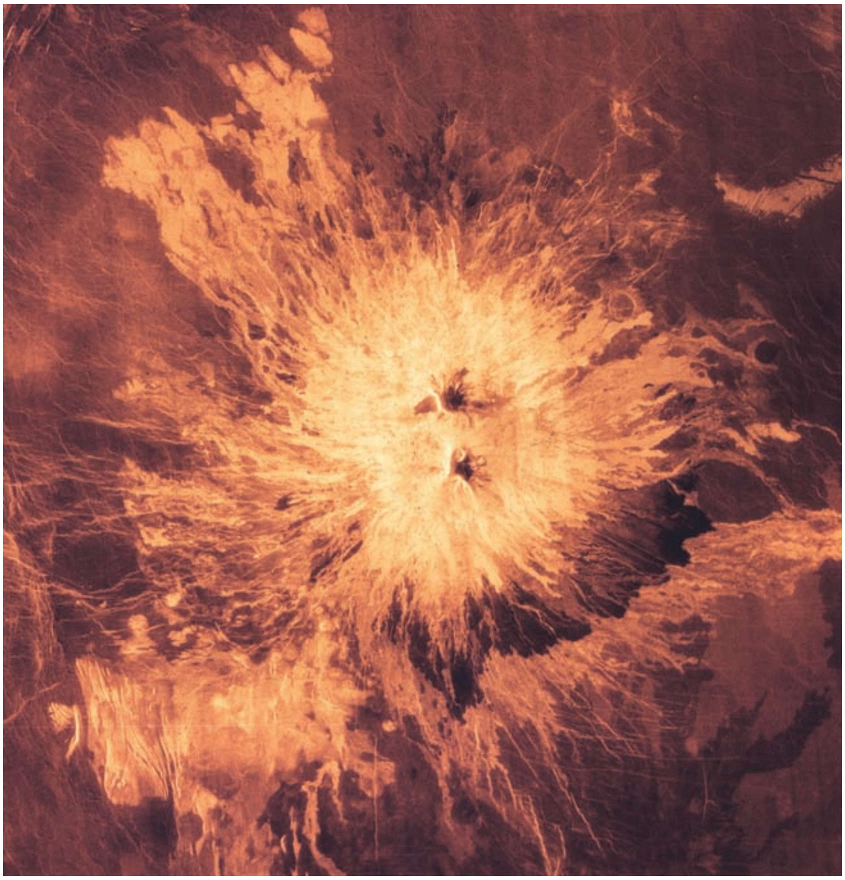


Fig. 7.16 Sapas Mons Located in an equatorial highland called Atla Regio, the volcano Sapas Mons is about 400 kilometers in diameter and 1.5 kilometers high; it is named after a Phoenician goddess. The sides of the volcano are composed of numerous overlapping lava flows from flank eruptions similar to terrestrial volcanoes such as the Hawaiian shield volcanoes. The summit contains two smooth, radar-dark mesas, as well as groups of pits, some as large as 1 kilometer across. They may have formed when underground chambers of magma were drained, causing the surface to collapse above them. A 20-kilometer-diameter impact crater northeast (*upper right*) of the volcano is partly buried in lava flows. (Courtesy of NASA/JPL.)

Aphrodite Terra, including Maat Mons and Sapas Mons (Fig. 7.16).

One of the highest volcanoes, Maat Mons, rises 9 kilometers above the surface, and spreads 200 kilometers across it (see Fig. 2.30). Sapas Mons is shorter and broader (Fig. 7.16). Both peaks are known as shield volcanoes, because they have the shape of a shield or an inverted plate. Similar giant shield volcanoes are found in the Hawaiian Islands, each with a broad base and gentle slopes.

Scientists suspect that some volcanoes on Venus may still be erupting. Infrared observations from the *Venus Express* spacecraft have provided evidence of minerals which were formed in relatively recent volcanic flows that occurred a few hundred to 2.5 million years ago. They suggest that the planet may still be geologically active. The recurrent volcanic release of sulfur dioxide and water vapor could also help replenish sulfur dioxide in the atmosphere of Venus.

The region of Venus that most closely resembles terrestrial mountains is Ishtar Terra, located at far northern latitudes. It consists of an elevated plateau, Lakshmi Planum, which is bounded on three sides by mountain belts – the Danu, Akna and Freyja Montes (Fig. 7.17). Lakshmi just drops off into the surrounding plains on the fourth side, forming an immense cliff. The belts of mountains, with their banded ridges and narrow valleys, resemble mountain ranges on Earth, and the loftiest, Maxwell Montes, rises to Himalayan altitudes, standing over 11 kilometers above the surrounding terrain.

The raised plateau and surrounding mountains closely resemble the Tibetan Plateau and the Himalayan Mountains on Earth, which were produced by the collision of India with Asia. However, since Venus has no colliding continents, the intensely deformed mountains must have been pushed up into the sky by a different process. One possible explanation suggests that the material beneath the mountains was cooler than surrounding areas and sank. The crust would then be pulled together, bunching up into mountains and compressing the surface features. Or perhaps that part of the crust was compressed and folded after the surrounding plains were formed, like a carpet pushed against a wall. In either case, Ishtar Terra is a unique type of raised feature on Venus, and different processes hold up other highland regions on the planet.

Gravity's highs and lows

How can the high places on Venus stay up when its rock temperatures are halfway to their melting point? Something has to be holding the mountains up. So

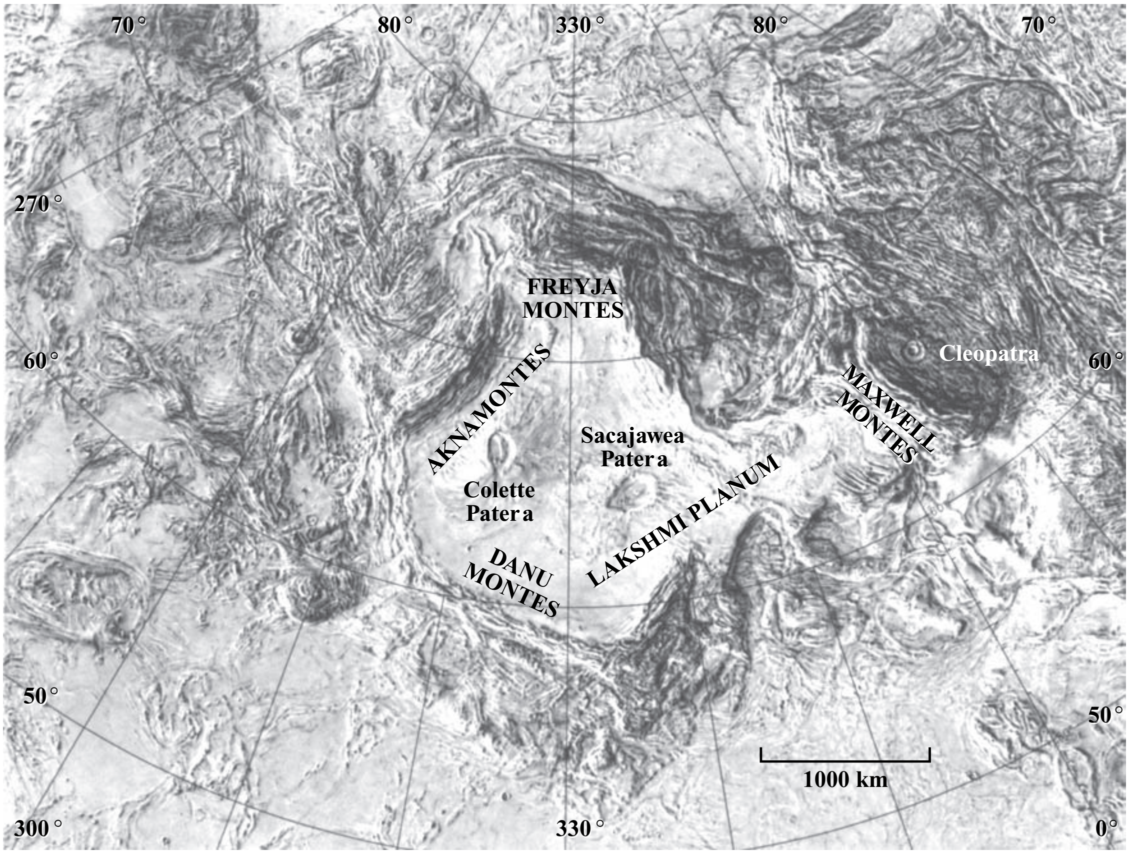


Fig. 7.17 Mountain ranges on Venus The Earth and Venus are the only planets in our solar system that have mountain belts. The highland massif of Ishtar Terra in the northern hemisphere of Venus includes a huge plateau, named Lakshmi Planum. It is about the size of Africa, rises 3.5 kilometers above the surrounding terrain, and is bordered by the Akna, Danu, Freyja and Maxwell mountains, each about 1000 kilometers in extent. Maxwell Montes stands 11 kilometers above the mean radius. Akna and Freyja are the respective names of the Yucatan goddess of birth and a goddess in Norse mythology. Danu is the Celtic mother of god. Maxwell is named after the British physicist James Clerk Maxwell, while Cleopatra is the Egyptian queen who had affairs with Julius Caesar and Mark Anthony. The Blackfoot Indian woman Sacajawea guided the Lewis and Clark expedition to the Pacific Northwest, and Colette (1873–1959) was a French novelist. (Adapted from a USGS map using Arecibo, *Pioneer Venus* and *Veneras 15* and *16* radar data.)

spacecraft measured local variations in gravity and used them to look underneath the highlands and see what is there.

The movements of the orbiting *Magellan* spacecraft were tracked by recording small changes in the wavelength of its radio signal, and the corresponding small changes in the orbital speed, inferred from the Doppler effect, were used to specify local variations in the underlying mass and density.

It turned out that some of the highest places on Venus exert the strongest gravitational pull, a correlation first noticed by tracking *Pioneer Venus Orbiter* and confirmed by sending *Magellan* into a low-altitude orbit. These volcanic rises, such as Atla Regio and Beta Regio, are probably held up by active motions below, fed by rising plumes of sluggish, upwardly buoyant flow. The hot, low-density material wells up from deep within the planet, with glacial slowness in spite of its heat, eventually spreading out beneath the surface, and pushing the volcanic rises up. Their low-density roots extend down under the high-standing regions and balance the mass excess. The combination is in "isostatic" equilibrium, like an iceberg floating on the ocean.

As the plumes near the surface, the lower pressures induce partial melting of the plume that sometimes punches holes in the crust, providing a conduit that permits lava to flow out from volcanoes and fissures in the planet. In this respect, the volcanic rises on Venus are thought to be very similar in origin to "hot spots" on Earth, such as the Hawaiian Islands.

7.7 Tectonics on Venus

Trapped heat deforms the surface

All planets have heat to get rid of. Gravity compressed and heated the planet's interiors when they formed, and

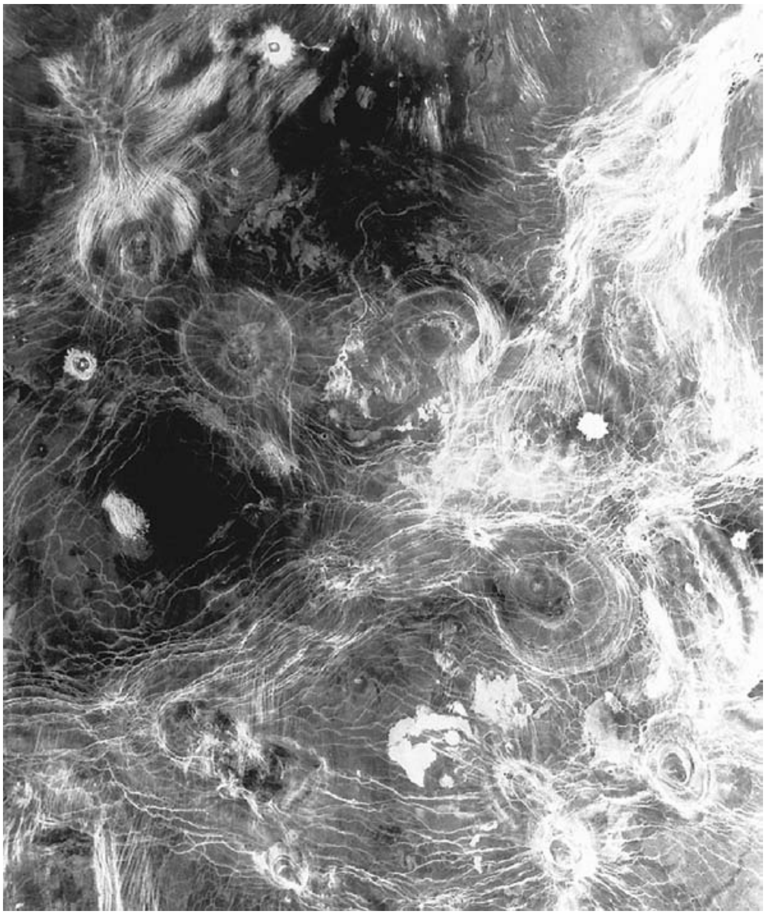


Fig. 7.18 Bereghinya Planitia This high-resolution *Magellan* mosaic shows an area in the low-lying plain Bereghinya Planitia, named for the Slavic water spirit. The image is 1840 kilometers wide, and centered at 45 degrees north latitude and 11 degrees east longitude. The plain's fractured surface is attributed to plumes of magma that rise from inside Venus, pushing against the planet's crust. The most prominent features are the circular and oval structures that are sometimes called arachnoids for the spider-web-like appearance of their fractures (also see Fig. 7.20). Also visible are lava flows, impact craters and volcanic domes. (Courtesy of NASA.)

colliding bodies also brought heat to them in their early stages. Then the decay of radioactive elements added more internal heat. It is still being generated within Venus by radioactive decay, producing molten rock inside the planet.

Size is the main factor in determining how much heat remains inside a planet, for larger bodies lose heat more slowly and remain internally active for longer times. Thus, the Earth's Moon, which is relatively small when compared to a planet, has not been volcanically active for 2 or 3 billion years, even though lava flowed into its impact basins to form the lunar maria before then. The larger, rocky planet Venus must have had vast churning reservoirs of hot material beneath its crust for a much longer period, and probably still does.

The heat trapped inside a planet wants to get out, so it rises to crack and deform the planet's surface or spills out in volcanoes. In technical terms, the molten rocks become swollen by heat and lower in density, rising through the cooler, overlying high-density material and carrying heat upward, like the convective bubbles in a pot of boiling water. The upward-flowing material wants to crack open or puncture a hole in the overlying material, breaking on through to the other side to release all that heat. The crustal deformations caused by the pent-up heat, and the internal changes affecting them, are known as *tectonics*, the Greek word for "building".

The hot rising material has buckled, fractured and stretched the crust on Venus, like a crumpled piece of paper or a face seamed and thickened by age (Fig. 7.18). It has split the crust open and spread it apart, forming rift valleys with steep sides and sunken floors. Some of them are found in Atla region of Venus alongside its volcanoes (Fig. 7.19). The linear rift zones in the equatorial highlands can extend for thousands of kilometers, but are cracked open by just a few kilometers. In contrast, rifts that split open the Earth's continents can, because of moving plates, widen up to make way for its biggest oceans.

When a bubble of hot material rises to just below the surface of Venus, it presses against the crust, causing the ground to bulge and crack. Circular and radial fractures are created around the edges of the rising dome, forming a network of radar-bright features that remind us of a spider's web (Fig. 7.20). Some of them have therefore been nicknamed *arachnoids*, from the Greek word for "spider". The term *coronae*, the Latin word for "crown", is used for the larger, elevated, circular structures that are also pushed up from below by rising molten rock trying to get out. Both

Fig. 7.19 Atla Regio Fractures, seen as bright, thin lines, criss-cross the volcanic deposits in part of the Atla region of Venus. The fractures are not buried by the lava, indicating that the convulsive tectonic activity post-dates most of the volcanic activity. This *Magellan* radar image is approximately 350 kilometers across, and centered at 9 degrees south latitude and 199 degrees east longitude. Several circular volcanoes, surrounded by radar-bright lobes, are also present. This region is named Atla after a Norse giantess, mother of Heimdall. (Courtesy of NASA/JPL.)

arachnoids and coronae are unique to Venus and have not been found on any other planet.

Coronae have concentric ridges and fractures that are hundreds of kilometers across, and large volcanic outpourings have occurred within them (Fig. 7.21). When enough lava spills out into a corona, the upwelling subsides and it is no longer supported from below. The bulge will deflate and buckle the surrounding terrain, producing an annulus of ridges and troughs that often surrounds coronae, like the moat around a castle. Or else, the magma cools and retreats as it ages and the molten rock drains back down the vent from whence it came. Then the dome will collapse like a giant fallen soufflé, creating ring-like fractures and a crumpled, cracked surface.

The increasing pressure of the upwelling magma can stretch the planet's skin until it bursts, like the broken cheese bubbles in a pizza or a split in an overcooked hotdog. Small volcanic domes, known as pancakes, are sometimes formed when pasty, sluggish lava breaks through and flows along the surface like toothpaste (Fig. 7.22). In other places the crust breaks and spreads open and lava flows into the gap like olive oil.

Just about everywhere that *Magellan*'s radar instrument looked, it found intersecting ridges, cracks and grooves, readily visible in the absence of overlying soil or vegetation and the absence of erosion. The intricate, tortured networks are found over the entire globe, ubiquitous throughout the volcanic plains and within the highlands. Internal forces associated with the pent-up heat have been pushing and pulling the crust, producing these patterns. They are known as *tesserae*, the Latin word for "tiles".

As the surface moves up in some locations and down in others, the associated stresses pull the surface apart or push it together. Over time, these stresses can be created in different directions, producing a regularly spaced grid pattern of fractured terrain that is only found on Venus at this scale (Fig. 7.23). Some of the cracked patterns of the tessarae have regular six-sided shapes that can be attributed to global heating and cooling of the surface. Repeated episodes of surface deformation in some highlands have additionally created a chaotic network of ridges, troughs and depressions with linear and curved structures.

Up, down and sideways on Earth and Venus

Radioactive elements decaying inside both the Earth and Venus generate enough heat to drive internal convection, in which hot material rises and cold material sinks on timescales of millions of years. The Earth's crust and uppermost mantle, its lithosphere, is broken into huge plates, each thousands of kilometers across, and convective

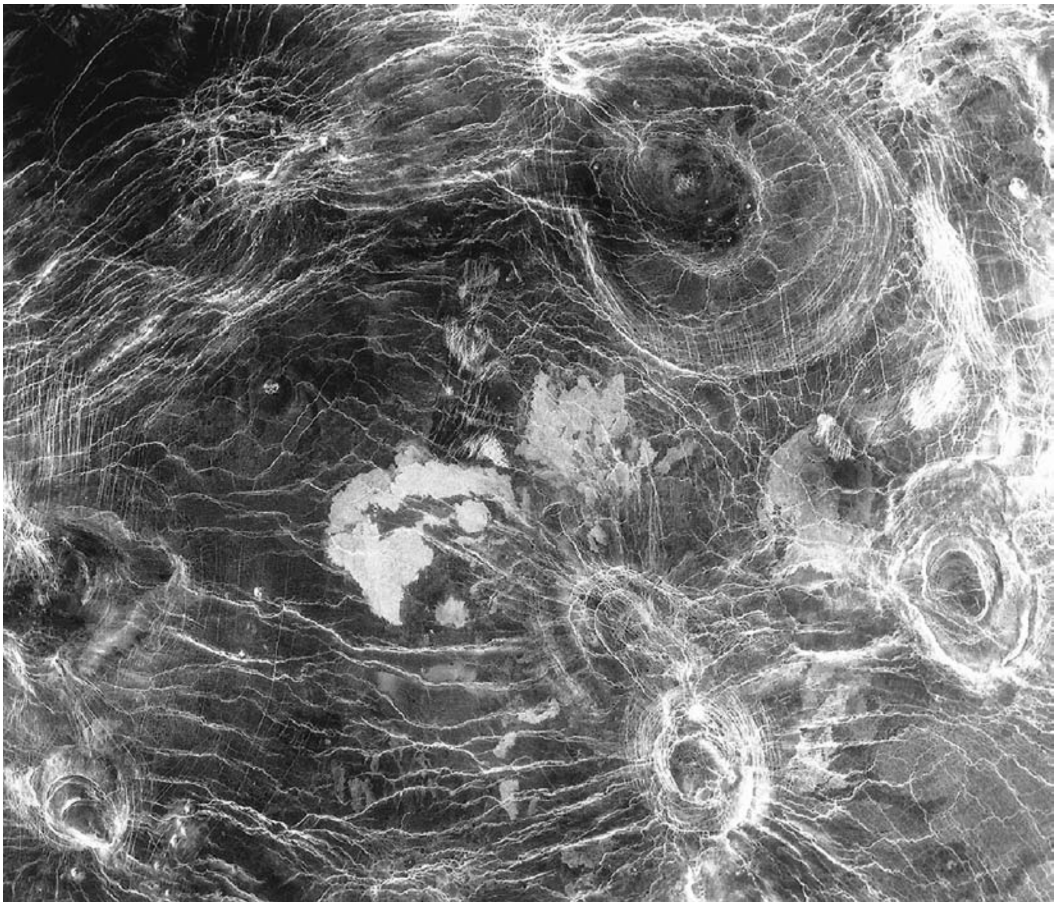


Fig. 7.20 Arachnoids An enlarged view of a *Magellan* mosaic of Bereghinya Planitia (also see Fig. 7.18), showing circular structures of up to 230 kilometers in diameter. Their central domes are surrounded by concentric and radial fractures. Such features have been informally dubbed arachnoids for their spider-web-like appearance. They are similar in form, but usually smaller than, the circular volcanic structures known as coronae, shown in Fig. 7.21. (Courtesy of NASA.)

currents propel the plates laterally, or sideways, across the globe. On Venus there is no trace of such a process.

Plate tectonics dominate the Earth's geology. The plates separate at mid-ocean ridges, descend into deep-ocean trenches where long arcs of volcanoes are made, collide with each other to form the great mountain ranges, and grind horizontally together to set off earthquakes. The plates are rigid, so any movement on one side also happens at the other; in between the plates move sideways, carrying the continents with them.

The Earth's crust is also recycled laterally. New ocean floor rises out of the interior at the mid-ocean ridges, moves horizontally for great distances across the planet, and is destroyed by diving into the deep-ocean trenches.

Because Venus and Earth are about the same size, and have a similar bulk composition, the rates of internal heat generation by radioactive decay and the energy available to drive internal motions should also be similar. Thus, there ought to be enough heat trapped inside Venus to push plates around its surface.

Yet there is no convincing evidence for a system of plates that slide horizontally about the surface of Venus, as they do on Earth. On Venus there are no features comparable to the Earth's extensive mid-ocean ridges or to its deep-ocean trenches. Thus, it is unlikely that the crust of Venus is recycled in the same way as the Earth's is.

Perhaps the reason Venus has no sliding plates is related to the lack of water on the planet. There can be no liquid water on the torrid surface, and the planet may also be exceptionally dry inside. Water in the Earth's lithosphere is believed to be essential to making it weak enough to break into plates in response to the motions of convection in the interior. On Venus, the present-day lithosphere is too strong to allow plates to develop. The outer shell of the planet is probably seized up tight, like a car engine without oil. But this may not have been true in the past when greater amounts of internal heat might have overcome the lack of lubricating water.

Whatever the exact explanation, the outer solid shell of Venus seems to consist of one thick, rigid plate, not many

Fig. 7.21 Fotla Corona Large circular and oval structures with diameters of 200 to 1000 kilometers are called coronae. The one shown at the center of this *Magellan* image is known as Fotla Corona, about 200 kilometers across and named after the Celtic fertility goddess. It is located in a vast plain to the south of Aphrodite Terra, and is centered at 59 degrees south latitude and 164 degrees east longitude. Molten rock rising from the interior of the planet most likely explains the corona's circular shape, raised topography, complex fractures and associated volcanism. Just north (*top*) of this corona is a flat-topped pancake dome, about 35 kilometers in diameter, thought to have formed by the eruption of sluggish, pasty lava. Another pancake dome is located inside the western (*left*) part of the corona. There is also a smooth, flat region in the center of the corona, probably a relatively young lava flow. Complex fracture patterns like the one in the northeast (*top right*) of the image are often observed in association with coronae. (Courtesy of NASA/JPL.)

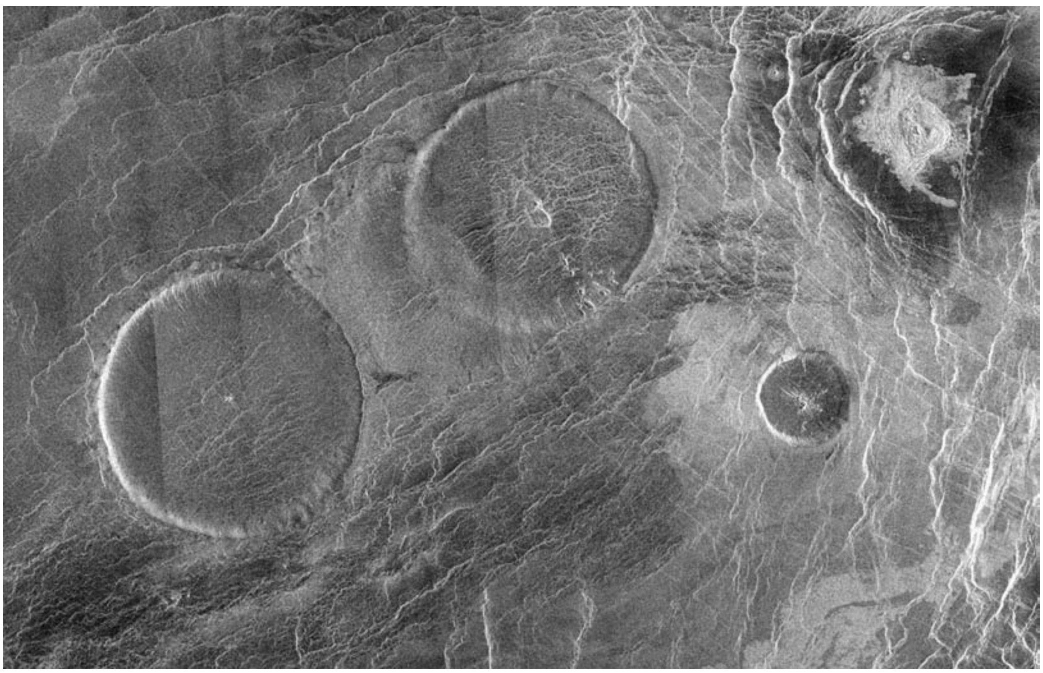


Fig. 7.22 Volcanic pancake domes This *Magellan* radar image, centered in the Eistia Regio at 12.3 degrees north latitude and 8.6 degrees east longitude, shows an area 250 kilometers wide. The prominent circular features are volcanic pancake domes, each about 65 kilometers across, with broad, flat tops less than one kilometer in height. They are formed from viscous, or sticky, lava, and include cracks and pits that result from cooling and withdrawal of the lava. (Courtesy of NASA/JPL.)

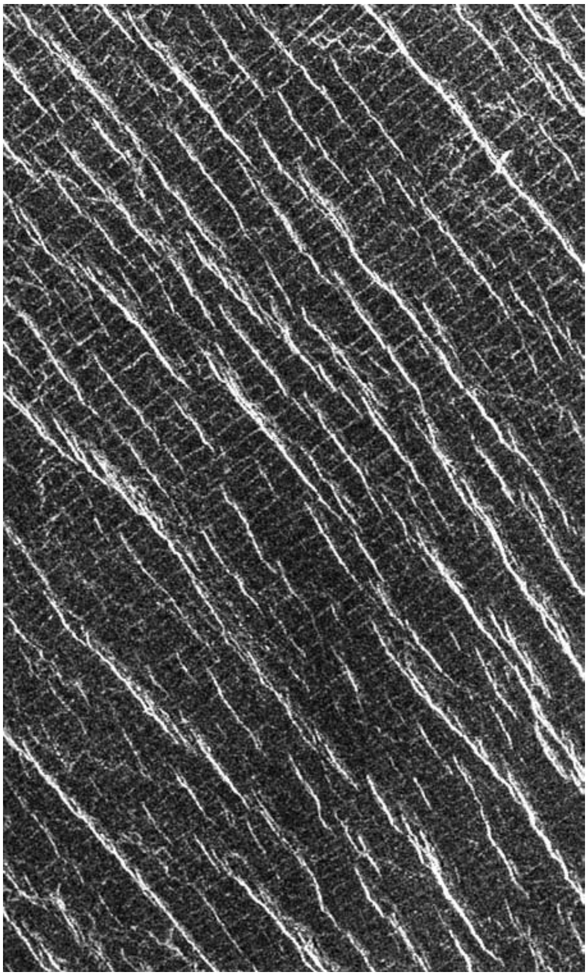


Fig. 7.23 Tesserae Because there is little soil and no vegetation or erosion to confuse us, we can see tectonic patterns on Venus much more easily than on our own planet. This region, covering an area of 37 kilometers wide by 80 kilometers long, is located at 30 degrees north latitude and 333 degrees east longitude, on the low rise separating Sedna Planitia and Guinevere Planitia. It shows a criss-crossed pattern of radar-bright lines, which appear to be faults or fractures. The orthogonal system of ridges and grooves is formed in elevated terrain (by 1 or 2 kilometers), and spaced at regular intervals of 1 to 20 kilometers. Known as *tessarae*, from the Latin word for "tiles", the features suggest repeated episodes of intense surface fracturing that may be unrelated to volcanic activity. This type of terrain is not seen on any other planet. It covers about 8 percent of Venus's surface. Sedna is an Eskimo goddess whose fingers became seals and whales, and Guinevere is the legendary Queen of the British King Arthur. (Courtesy of NASA.)

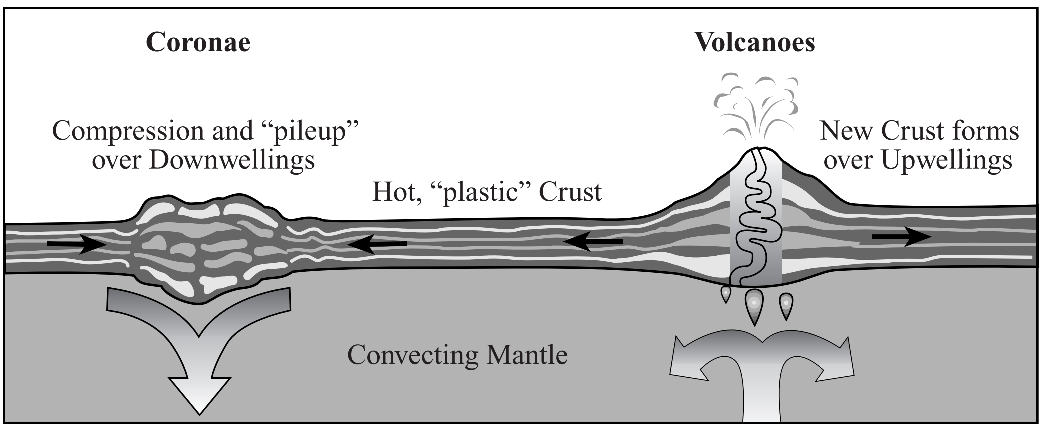


Fig. 7.24 Vertical motions on Venus Some of the surface features on Venus have been formed by vertical, up-and-down motions driven by hot material welling up from the planet's interior. When rising bubbles of hot rock press against the crust, they can form domed, cracked features known as coronae. Larger plumes support volcanic rises on Venus, while sinking regions can lead to mountain formation.

shifting plates as on Earth. Such a strong and unbroken lithosphere would also help support the high plateaus on Venus over long timescales, and it might also explain why most of the surface is at one low level.

Given the absence of plate tectonics, there is only one main way for Venus to lose its heat. The hot mantle material pushes up and the dominant movement is often vertical, or up and down (Fig. 7.24). Upwelling material pushes

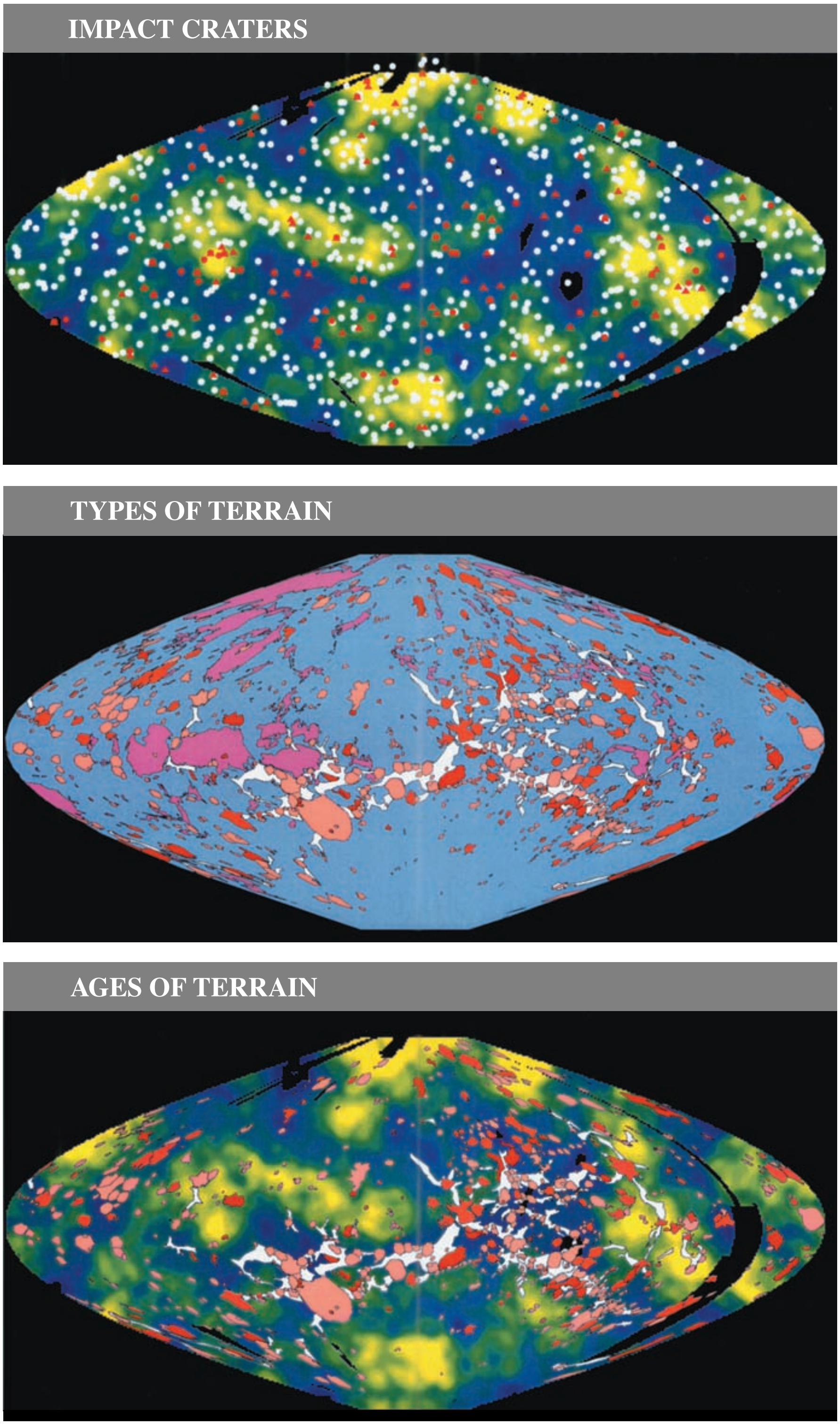


Fig. 7.25 Impact craters, terrain type, and terrain ages on Venus Impact craters (*top diagram*) are randomly scattered all over Venus. Most are pristine (*white dots*). Those modified by lava (*red dots*) or by faults (*triangles*) are concentrated in places such as Aphrodite Terra. Areas with a low density of craters (*blue background*) are often located in highlands. Higher crater densities (*yellow background*) are usually found in the lowland plains.

The terrain type (*middle diagram*) is predominately volcanic plain (*blue*). Within the plains are deformed areas such as tessarae (*pink*) and rift zones (*white*), as well as volcanic features such as coronae (*peach*), lava floods (*red*) and volcanoes of various sizes (*orange*). Volcanoes are not concentrated in chains as they are on Earth, indicating that plate tectonics do not operate.

Terrain age data (*bottom diagram*) indicate that volcanoes and coronae tend to clump along equatorial rift zones, which are younger (*blue*) than the rest of the surface of Venus (*green*). Tesserae, ridges and plains are older (*yellow*). In general, however, the surface lacks the extreme variation in age that is found on Earth and Mars. (Courtesy of NASA/Mary Beth Price.)

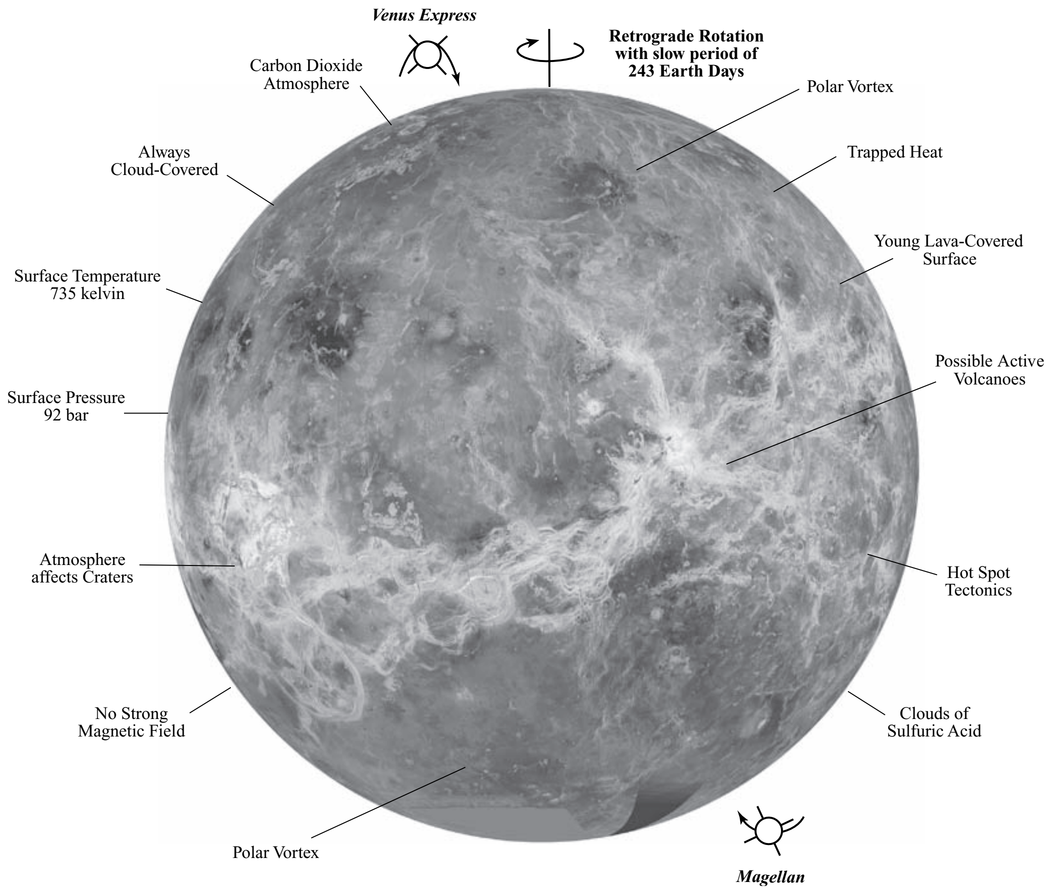


Fig. 7.26 Summary diagram

against the ground, creating arachnoids, coronae and tessarae, and punctures the surface to form volcanoes. Volcanic rises are held up by the hot rising material, and long mountain ranges may have been built during sinking, downward compression. Moreover, vast regions of the planet consist of flat, lowland plains with no substantial motion, either vertically or horizontally.

Although astronomers know virtually nothing about the first 4 billion years on Venus, they have been able to piece together a sequence of events during the last 750 million years (Fig. 7.25). At the beginning of the record, they see isolated surface deformation giving rise to a locally fractured crust. Widespread lava flooding created the flat lowland plains soon after this episode of tessera formation. After this brief but intense period of global volcanic floods, the style and rate of volcanism changed. Localized volcanoes grew on top of the vast plains and coronae were formed within them, but primarily in the equatorial regions where extensive rifts are also found.

8 Mars: the red planet

- Like the Earth, the planet Mars has an atmosphere, white clouds, polar caps and seasons.
- Mars has a partially liquid core, probably containing molten iron and perhaps surrounding a solid iron core, as within the Earth.
- Mars does not now have a global dipolar magnetic field to deflect lethal cosmic rays and energetic solar particles.
- The oldest terrain on Mars exhibits bands of magnetized material with alternating polarity, most likely originating about 4 billion years ago when the red planet might have had a global dipolar magnetic field.
- In the early 20th century it was thought that seasonal water melting from the polar caps in spring and summer produced a dark band of vegetation on Mars, and that intelligent Martian inhabitants had constructed canals to transport water across the planet.
- The seasonal dark regions on Mars are now attributed to winds, and the canals are now known to be an illusion caused when the eye arranges small, disconnected features into lines.
- Mars now has a thin, cold and dry atmosphere that is composed almost entirely of carbon dioxide, with a surface pressure of about a one-hundred-fiftieth of the Earth's atmosphere and a mean surface temperature that is well below the freezing temperature of water.
- Because of the low surface pressure and temperature of the Martian atmosphere, it cannot now rain on Mars. If any liquid water were now released on the planet's surface, it would survive for just a brief time before freezing or evaporating.
- The Martian atmosphere contains virtually no oxygen, so it has no ozone layer. The planet's surface is therefore exposed to the full intensity of the Sun's ultraviolet rays.
- Powerful and pervasive winds roar across Mars, sweeping up vast dunes of sand and fine-grained dust, creating tornado-like dust-devils, and occasionally producing global dust-storms that hide the entire planet from view.
- Seasonal polar caps of frozen carbon dioxide, or dry ice, wax and wane with the seasons on Mars. They lie on top of extensive caps of frozen water in both hemispheres. In the north, a residual or permanent cap of water ice remains in the summer heat, but carbon dioxide ice also remains in the southern winter.

- Layers in the polar caps of Mars suggest climate changes on timescales of 10 000 to 100 000 years, perhaps triggered by periodic variations in the planet's orbit and rotation axis.
- Mars is divided into two strikingly different hemispheres. In the south there are the older, elevated, heavily cratered highlands that resemble the lunar highlands. In the north there are the younger, lower-lying, smoother volcanic plains.
- Towering volcanoes and immense canyons are found on Mars.
- The dry tracks of past flowing water are etched into the surface of Mars, marking the site of ancient rivers and floods that occurred 3 to 4 billion years ago. Water networks are found in the heavily cratered southern highlands, and outflow channels are located in the equatorial regions running downhill from the highlands to the northern plains.
- Water might have once lapped the shores of long-vanished lakes and seas on Mars.
- Instruments aboard orbiting spacecraft have detected water-related minerals on the surface of Mars, especially in the ancient southern highlands.
- The *Mars Pathfinder* lander, with its *Sojourner Rover*, found that its landing site has been untouched by water since it flowed across the region more than 2 billion years ago. The two *Mars Exploration Rovers*, *Opportunity* and *Spirit*, have found evidence of ancient water flow on Mars, but no signs of recent running water.
- Huge amounts of water once flowed on the Martian surface, but exactly where all that water came from and where it all went is still uncertain. Colliding asteroids and comets could have deposited the water in the early history of the planet, or Mars might have been warmer long ago, with a thick, dense atmosphere, rain and flowing water. Most of the water that once flowed on Mars is now frozen into ice on or below the surface.
- Subsurface water ice is suggested by flow-like patterns of material ejected from impact craters on Mars, and substantial amounts of frozen water have been inferred from spectrometers aboard the *2001 Mars Odyssey* spacecraft.
- The *Phoenix* lander obtained evidence for subsurface water ice and past water-flow in the northern polar plains of Mars, and detected snow falling from Martian clouds of water ice.
- Liquid water may have been seeping out of the walls of canyons and craters on Mars in recent years, creating small gullies and depositing the debris in fanlike deltas. Other gullies have been attributed to landslides of loose dust or flows of sand related to carbon-dioxide frost.
- The lack of a global dipolar magnetic field and a thick atmosphere with an ozone layer, which would respectively divert cosmic rays or solar energetic particles and absorb ultraviolet rays, suggest that Mars' surface is now a sterile place where life cannot survive. If life did once exist there, it might have survived beneath the surface, within rocks or deep underground in the possibly wet and more temperate part of the planet's interior.
- Five spacecraft have successfully landed on the reddish-brown Martian surface and revealed no signs of life on Mars.
- There are no detectable organic molecules or cells in the Martian surface examined by the *Viking 1* and *2* landers. They found no unambiguous evidence for biological life at their landing sites. The chemically reactive, highly oxidized soil will destroy cells, while also rusting the Martian surface red.

- Cosmic impacts with Mars are capable of ejecting surface rocks into space, and some of them eventually arrive at the Earth. One such meteorite from Mars, named ALH 84001, was once thought to exhibit evidence for ancient, microscopic bacteria-like fossils, but non-biological explanations are now accepted for these features.
- The search for life on Mars is now focused on the hunt for liquid water, which might indicate that Mars is habitable and could have sustained life either in the remote past or recently.
- Methane has recently been found in the Martian atmosphere; it could be due to geochemical processes or to bacteria-like microorganisms.
- NASA has plans to launch a *Mars Science Laboratory* that would analyze the surface material near its landing site on Mars for proteins, amino acids and other molecules that are essential to life as we know it.
- Mars has two small moons, named Phobos and Deimos. Phobos is heading towards eventual collision with Mars.

8.1 Fundamentals

The red planet Mars is only half the size of Earth, and its mass is one-tenth of the Earth's mass. Mars has a day that is just 37 minutes longer than the Earth's day, and the length of the year on Mars is about twice that of a year on Earth. Mars exhibits clouds, variable polar caps and seasons, and its freezing surface lies beneath a cold, thin atmosphere that is composed mainly of carbon dioxide.

Table 8.1 Physical properties of Mars[a]

Mass	6.4169×10^{23} kilograms $= 0.1074\ M_E$
Mean radius	3.3895×10^{6} meters $= 0.532\ R_E$
Mean mass density	3933.5 kilograms per cubic meter
Surface area	1.441×10^{14} square meters $= 0.2825\ A_E$
Rotation period or length of sidereal day	24 hours 37 minutes 22.663 seconds $= 8.86427 \times 10^{4}$ seconds $= 1.02596$ Earth days
Orbital period	1.880 847 6 Earth years = 686.98 Earth days = 668.60 Mars solar days
Mean distance from Sun	2.2794×10^{11} meters = 1.523 66 AU
Orbital eccentricity	0.0934
Tilt of rotational axis, the obliquity	25.19 degrees
Distance from Earth	5.6×10^{10} meters to 3.99×10^{11} meters
Angular diameter at closest approach	14 to 24 seconds of arc
Age	4.6×10^{9} years
Atmosphere	95.32 percent carbon dioxide, 2.7 percent nitrogen, 1.6 percent argon
Surface pressure	3 to 14 millibars = 0.003 to 0.014 bars
Surface temperature range	140 to 290 kelvin
Average surface temperature	210 to 220 kelvin
Magnetic field strength (remnant)	$\pm 1.5 \times 10^{-6}$ tesla $= \pm 0.05\ B_E$
Magnetic dipole moment	Less than 10^{-4} that of Earth

[a] Adapted from H. H. Kieffer, B. M. Jakosky, C. W. Snyder and M. S. Matthews (eds.), *Mars,* University of Arizona Press, Tucson, 1992. The symbols M_E, R_E, A_E, and B_E respectively denote the mass, radius, surface area and magnetic field strength of the Earth.

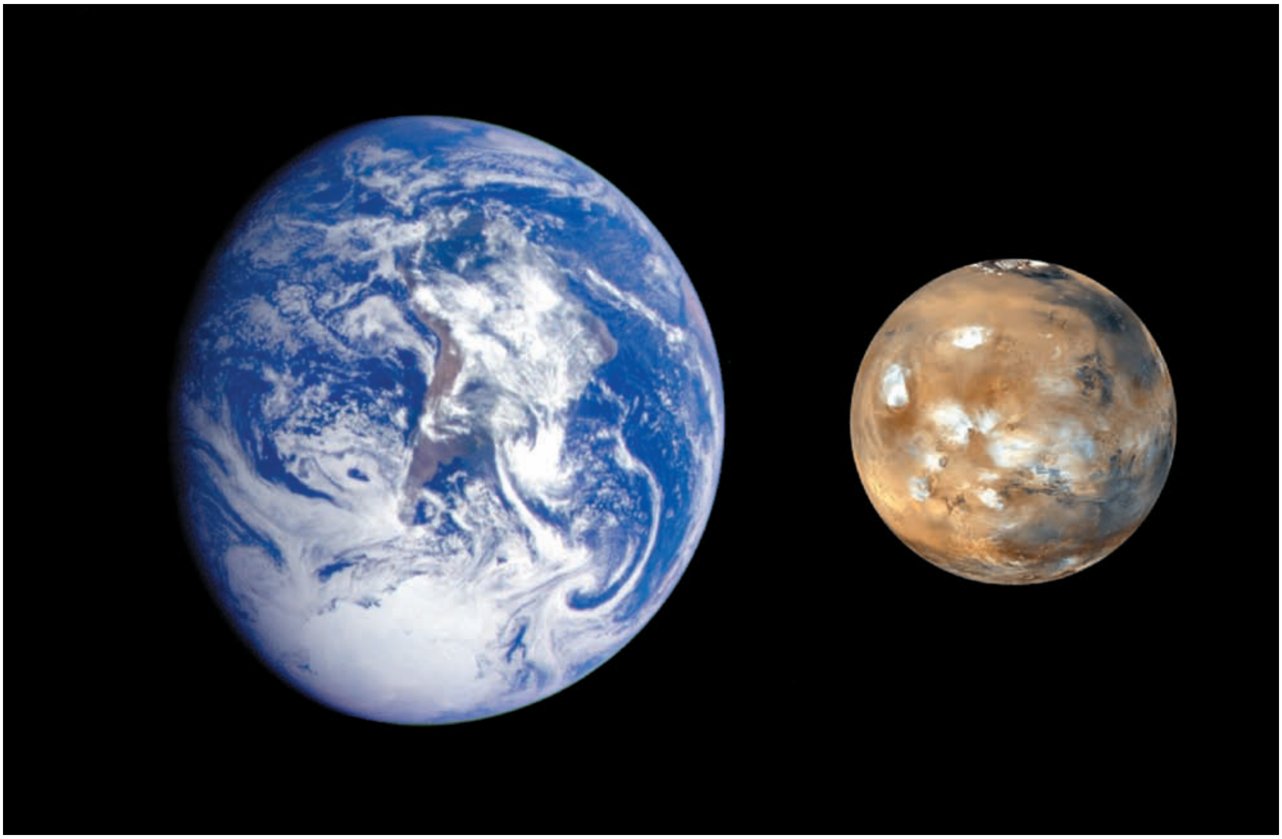


Fig. 8.1 Earth and Mars This composite image demonstrates the relative size, similarities and main difference of Earth (*left*) and Mars (*right*). Mars is about half the size of the Earth. Both planets exhibit clouds and polar caps. Bluish-white clouds of water ice hang above volcanoes on Mars (*right – center left*) and large dark areas extend across its red surface (*right – top right*). The residual north polar cap on Mars (*right top*) is made of water ice, and is circled by dark dunes of sand and dust. The Earth also has clouds and polar caps composed of water ice. However, about 75 percent of the Earth is covered with oceans, while liquid water cannot now exist for long times on the surface of Mars. The Earth image was taken from the *Galileo* spacecraft in December 1990, and the *Mars Global Surveyor* acquired the Mars image in April 1999. (Courtesy of NASA/JPL/MSSS.)

8.2 Planet Mars

Mars is an Earth-like planet

Mars, fourth planet from the Sun, is our closest planetary neighbor after Venus, and it ranks third in brightness as seen from Earth – after Venus and Jupiter. The red planet is brighter than most stars, and it has intrigued humans since prehistoric times because of its reddish color and slow looping movement across the starry background. The ancients associated the blood-red color with destruction and warfare. The Babylonians knew it as Nergal, or Nirgal, the star of death, and the Greeks and Romans named the wanderer after their gods of war – Ares and Mars, respectively.

In several ways, Mars is similar to the Earth. The red planet spins on its axis with a rate and tilt that are almost identical to the Earth's. Surface markings have long allowed astronomers to determine the planet's rotation period. It rotates once on its axis every 24.62 hours, so the day on Mars is only 37 minutes longer than our own. Because Mars' rotational axis is inclined to its orbital plane, Mars has seasons. In fact, the present tilt of Mars' axis, at 25.2 degrees, is similar to that of Earth, at 23.5 degrees. Both planets therefore have four seasons – autumn, winter, spring and summer – although the Martian seasons last about twice as long since the Martian year is 687 Earth days long or 1.9 Earth years. But Mars is a considerably smaller planet, about half the size and one-tenth the mass of Earth (Fig. 8.1).

The eccentricity of Mars' orbit is larger than Earth's, 0.093 compared with 0.015, so the shape of the orbit of Mars is more elliptical, or out of round, than the orbit of Earth, which is closer to a circular shape. As a consequence, differences between the northern and southern seasons are much more pronounced on Mars than on Earth. The Martian pole that tilts toward the Sun at Mars' closest approach to the star, or at perihelion, has warmer summers than the other pole, and at present the south has the warmer summers. That is, the closest approach of Mars to the Sun currently occurs during the southern summer and northern winter. But because of slow changes in the direction of tilt of the rotational axis and the orientation of the perihelion, the hot and cold poles change on a 51 000-year cycle.

The Earth-like appearance of Mars in a telescope has intrigued humanity for centuries. Both planets have an atmosphere, clouds, polar caps and seasons. An atmosphere is required for the formation of the white clouds, which are composed of water-ice crystals that freeze out of the high cold air, like clouds on Earth. Clouds of carbon-dioxide ice also form above the Martian water-ice clouds. The polar caps on Mars grow in the local winter, when gases are extracted from the atmosphere, and recede with the coming of local spring and summer when gases are released into the Martian atmosphere.

Mars' kinship with the Earth has recently been extended to their interiors. After years of tracking the radio signals sent from spacecraft that have landed on Mars or are orbiting the planet, scientists have measured a small tidal bulge in the solid body of Mars, towards and away from the Sun, and determined the density at various depths in Mars. Like the Earth, the red planet is not merely a solid ball of rock, but instead contains a dense, metallic core. Early in Mars' history, the molten rock on Mars became differentiated, with heavy elements like iron sinking to the center and the light ones rising to the top. The size of this core is large enough to indicate that it must be partially liquid, most likely containing iron that is at least partly molten. But the absence of any detectable global dipolar magnetic field and the relatively high iron abundance at the surface of Mars suggest that Mars never melted as extensively as did Earth.

Right now Mars lacks an internally generated global dipolar magnetic field, which the Earth has, so the resemblance stops there. Nevertheless, *Mars Global Surveyor* detected bands of residual magnetism with alternating magnetic polarity, embedded in the ancient Martian surface (Fig. 8.2). Their presence indicates that Mars had a magnetic field in the past, but that it switched off about 4 billion years ago. The leftover magnetic bands may have been caused by a global magnetic field that repeatedly switched polarity as new crust was forming, somewhat like the alternating magnetic polarities found in the spreading ocean floor on Earth.

Early speculations about intelligent life on Mars

Over the past century, our fascination with Mars has been stimulated largely by the prospect that life may exist there, either in the past or the present. Large, varying dark regions form on the planet, seasonally distorting its ruddy face (Figs. 8.3, 8.4). The grayish-green regions flourish in the summer and become dormant in winter, as many plants do on Earth. Their seasonal growth on Mars has been called the "wave of darkening" since a dark band moves from the southern polar cap toward the equator as the cap shrinks. It was once thought that water melting from the polar cap might cause hypothetical vegetation to grow and progress from the southern polar region to the equator.

In the early 20th century, it was even widely believed that advanced civilizations had developed on Mars. These speculations resulted from ground-based telescopic observations apparently showing oases and canals stretching across the dusty plains of Mars, but often glimpsed at the limit of telescopic visibility.

In 1877, when Mars was exceptionally close to the Earth, the Italian astronomer Giovanni Schiaparelli (1835–1910), director of the Milan Observatory, reported that a maze of dark, narrow straight lines traverses the planet's surface (Fig. 8.5). He called them *canali*, the Italian word for "channels", assuming that they were natural features. Schiaparelli mapped them and gave the broadest ones the names of large terrestrial rivers, such as the Ganges and Indus.

The French astronomer Camille Flammarion (1842–1925) subsequently wrote in 1892 that the channels resemble man-made canals, redistributing scarce water across a dying Martian world. Flammarion was also convinced that the Martian inhabitants might be more advanced than terrestrial humans, describing them in popular books.

At about the same time, a wealthy Bostonian named Percival Lowell (1855–1916) convinced much of the American public that there was intelligent life on Mars. Rich enough to do as he pleased, Lowell built an observatory in the clear air of Flagstaff, Arizona, with the specific intention of observing and explaining the Martian canals. When Lowell turned his telescope toward Mars in 1894, he found what he expected to see – a vast network of canals bordered by vegetation. He also published popular books attributing the canals to a vast, planet-wide irrigation network, constructed by intelligent beings to transport water away from the melting polar caps to parched equatorial deserts.

Most astronomers, however, could not see the canals, concluding that they were an optical illusion if they existed at all. As it turned out, the "canals" have no objective reality, beyond the tendency of the human mind to seek order in chaos.

On close inspection from spacecraft, the so-called canals dissolve into swarms of elongated light and dark streaks, often tens of kilometers long, pointing in the

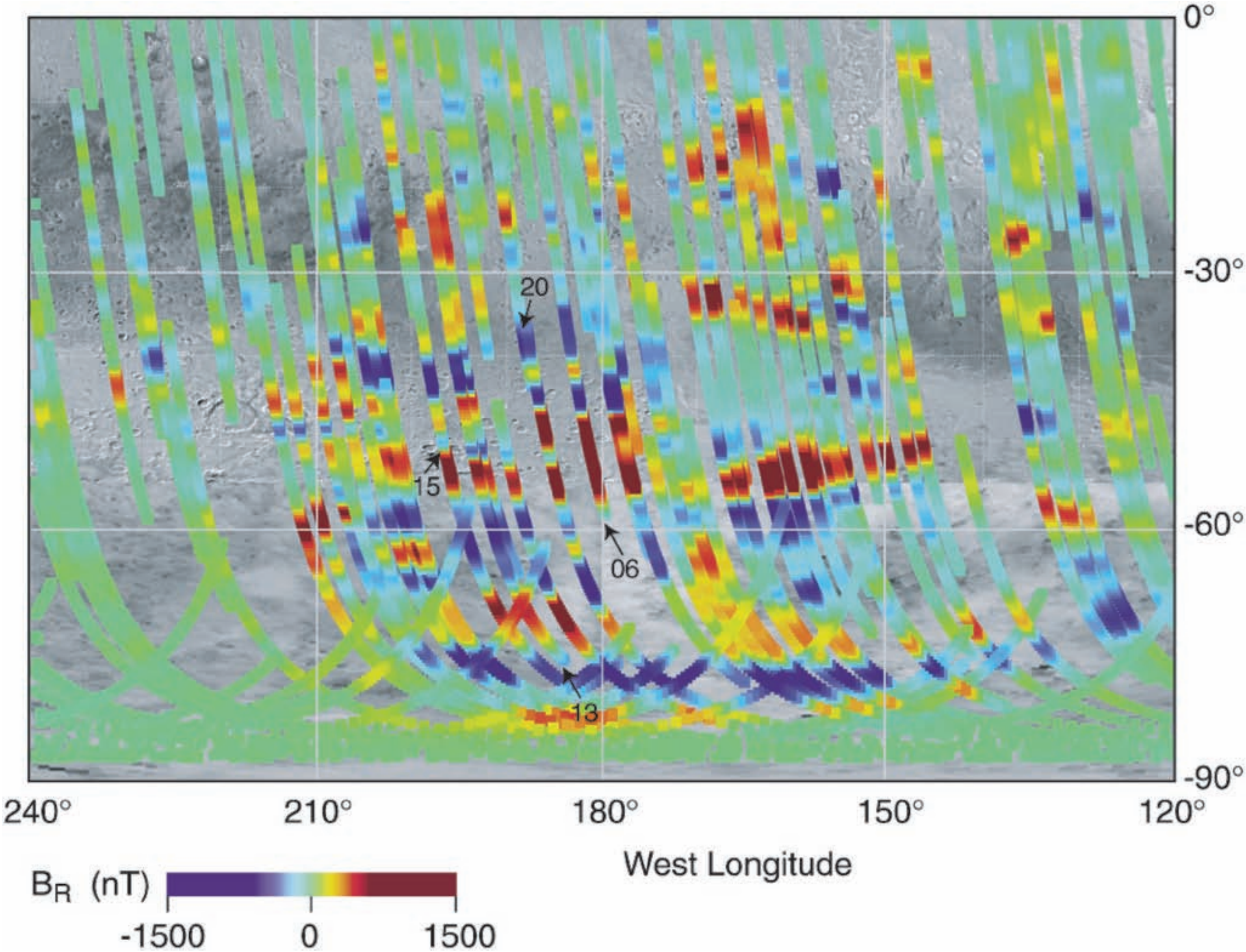


Fig. 8.2 Magnetic stripes on Mars Alternating bands of magnetic polarity are most prominent in this part of the southern highlands, near Terra Cimmeria and Terra Sirenium. The magnetic data were obtained from an instrument aboard the *Mars Global Surveyor*. This map is color-coded red for a positive magnetic field pointing out of the planet and blue for a negative one pointing in, with a strength up to 1500 nanotesla, or 1.5×10^{-6} tesla. Stripes of alternating polarity, or direction, extend up to 2000 kilometers across the planet in the east–west direction. They are similar to the magnetic patterns seen in the Earth's crust at both sides of the mid-oceanic ridges, where the spreading crust has recorded flip-flop reversals in the Earth's dipolar magnetic field. (Courtesy of NASA/ JPL/GSFC.)

directions of strong prevailing winds. When all the streaks in a given area are integrated and superimposed by the human eye when looking through a telescope, they form canal-like features, in much the same way as dots in newsprint combine to make a picture. Moreover, the wave of darkening is not a sign of the seasonal revival of life, but instead develops when the surface rocks are scoured by powerful, seasonal winds that remove fine dust deposited on the darker terrain. Thus, the basis for early speculations about intelligent life on Mars was illusory.

8.3 The space-age odyssey to Mars

Detailed ground-based telescopic observations of Mars are difficult. One reason is that the planet is out of view for prolonged periods when it is on the other side of the Sun from the Earth. The other reason is that turbulence in the terrestrial atmosphere limits the angular resolution of even the most powerful telescope on Earth.

When Mars is closest to us and most easily observed with ground-based telescopes, the Earth moves between Mars and the Sun and the planet Mars is fully illuminated by sunlight, casting few shadows and preventing the detection of topographic detail. This alignment occurs every 780 days, or 26 months, and is known as an opposition, since Mars is then opposite the Sun in our sky (Table 8.2). Even when Mars is in opposition, the biggest telescope on Earth can separate or resolve details no smaller than about one-twentieth of the Martian diameter or 300 kilometers across, providing only a blurred vision

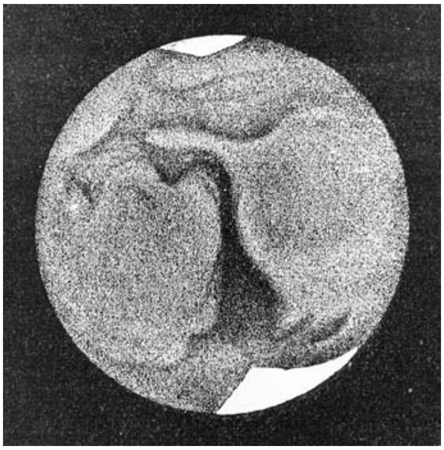


Fig. 8.3 Polar cap of Mars and its Syrtis Major The English amateur astronomer Warren De La Rue (1815–1889) made this drawing of Mars on 20 April 1856, using a 0.33-meter (13-inch) reflector telescope. It shows bright polar caps and a dark, triangular feature now known as Syrtis Major Planitia (Gulf of Sirte Plains). The Dutch astronomer Christiaan Huygens (1629–1695) first sketched this feature in 1659. From his observations of Syrtis Major, Huygens concluded that the rotation period of Mars is about 24 hours. This drawing is reproduced from Camille Flammarion's (1842–1925) book *La Planete Mars et ses Conditions d'Habitabilité*, published in 1892.

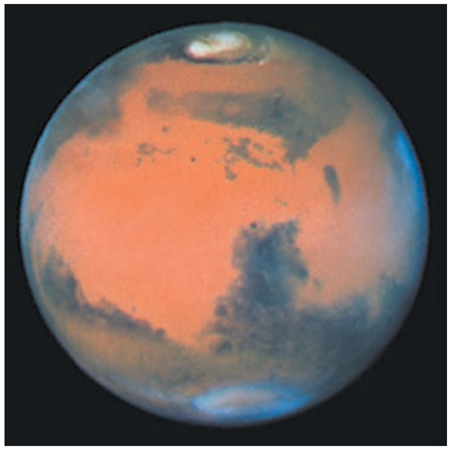


Fig. 8.4 *Hubble Space Telescope* views Mars This perspective of Mars was obtained from the *Hubble Space Telescope* (*HST*) on 10 March 1997 – the last day of spring in the Martian northern hemisphere. The red planet was near its closest approach to Earth and a single picture element of the *HST* spanned 22 kilometers on the Martian surface. The image shows bright and dark markings observed by astronomers for more than a century. The large dark feature seen just below the center of the disk is Syrtis Major Planitia, first seen telescopically by Christiaan Huygens (1629–1695) in the 17th century. To the south of Syrtis Major is the large circular impact basin Hellas (*center bottom*) filled with surface frost and shrouded in bright clouds of water ice. The seasonal north polar cap (*top center*) is rapidly sublimating, or evaporating from solid dry ice to carbon dioxide gas, revealing the smaller residual water ice cap with its collar of dark sand dunes. (Courtesy of NASA/JPL/STScI/David Crisp.)

of Mars. To see Mars in greater detail, one must use a spacecraft to approach it more closely.

The United States National Aeronautics and Space Administration (NASA) has dominated the modern exploration of Mars, sending numerous spacecraft to image the planet from orbit or to land on its surface. Systematic mapping of the Martian surface began in 1972, when NASA's *Mariner 9* orbiting spacecraft revealed the planet's volcanoes, deep canyons, huge outflow channels caused by ancient catastrophic floods, and extensive dune fields. A few years later, the two *Viking* orbiters fully mapped the surface with a resolution of about 250 meters, and the two *Viking* landers examined the chemistry of the local soil and carried out an unsuccessful search for signs of life on Mars.

In the late 20th and early 21st centuries, the NASA Mars Exploration Program has continued to seek evidence for whether Mars was, is, or can be a habitable world, using orbiters with high-resolution cameras, capable of resolving features a few meters across, and landers that directly sample the surface and its environment (Tables 8.3, 8.4). Since living things require water, the recent scientific strategy has also involved a search for water, either liquid or frozen and in the past or present (Sections 8.7, 8.8). The overall goals of the program are to determine if life ever arose on Mars, to characterize the climate and geology of the planet, and to prepare for human exploration of Mars.

8.4 The atmosphere, surface conditions and winds of Mars

A carbon-dioxide atmosphere

During the first half of the 20th century, astronomers guessed that the Martian atmosphere was mostly nitrogen, like that of the Earth. Carbon dioxide had been identified in the planet's atmosphere by mid century, but it was thought to be a minor ingredient. Then instruments aboard *Mariner 4* showed in 1965 that carbon dioxide is the primary gas, and nitrogen is a minor constituent, while also revealing an ancient, cratered terrain that resembles the lunar highlands.

The *Viking 1* and *2* landers made detailed measurements of the composition of the Martian atmosphere in 1976 (Table 8.5). Carbon dioxide is indeed the principal constituent, amounting to 95.32 percent of the atmosphere at ground level, followed by nitrogen (2.7 percent) and

Table 8.2 Oppositions of Mars 2001 to 2035[a]

Opposition date	Right ascension (hours minutes)		Declination (degrees minutes)		Diameter (seconds of arc)	Distance[b] (10^7 km)
2001 June 13	17	28	−26	30	20.5	6.82
2003 Aug. 28	22	38	−15	48	25.1	5.58
2005 Nov. 7	02	51	+15	53	19.8	7.03
2007 Dec. 28	06	12	+26	46	15.5	8.97
2010 Jan. 29	08	54	+22	09	14.0	9.93
2012 Mar. 3	11	52	+10	17	14.0	10.08
2014 Apr. 8	13	14	−05	08	15.1	9.29
2016 May 22	15	58	−21	39	18.4	7.61
2018 July 27	20	33	−25	30	24.1	5.77
2020 Oct. 13	01	22	+05	26	22.3	6.27
2022 Dec. 8	04	59	+25	00	16.9	8.23
2025 Jan. 16	07	56	+25	07	14.4	9.62
2027 Feb. 19	10	18	+15	23	13.8	10.14
2029 Mar. 25	12	23	+01	04	14.4	9.71
2031 May 4	14	46	−15	29	16.9	8.36
2033 June 27	18	30	−27	50	22.0	6.39
2035 Sept. 15	23	43	−08	01	24.5	5.71

[a] An opposition occurs when the Earth moves between Mars and the Sun, and the two planets are closest. Adapted from William Sheehan, *The Planet Mars*, University of Arizona Press, Tucson, 1996.

[b] The distance between the Earth and Mars at opposition in units of 10 million kilometers (10^7 km). Because the Martian orbit is more elliptical than Earth's, the distance between the two planets at different oppositions varies as much as 50 million kilometers.

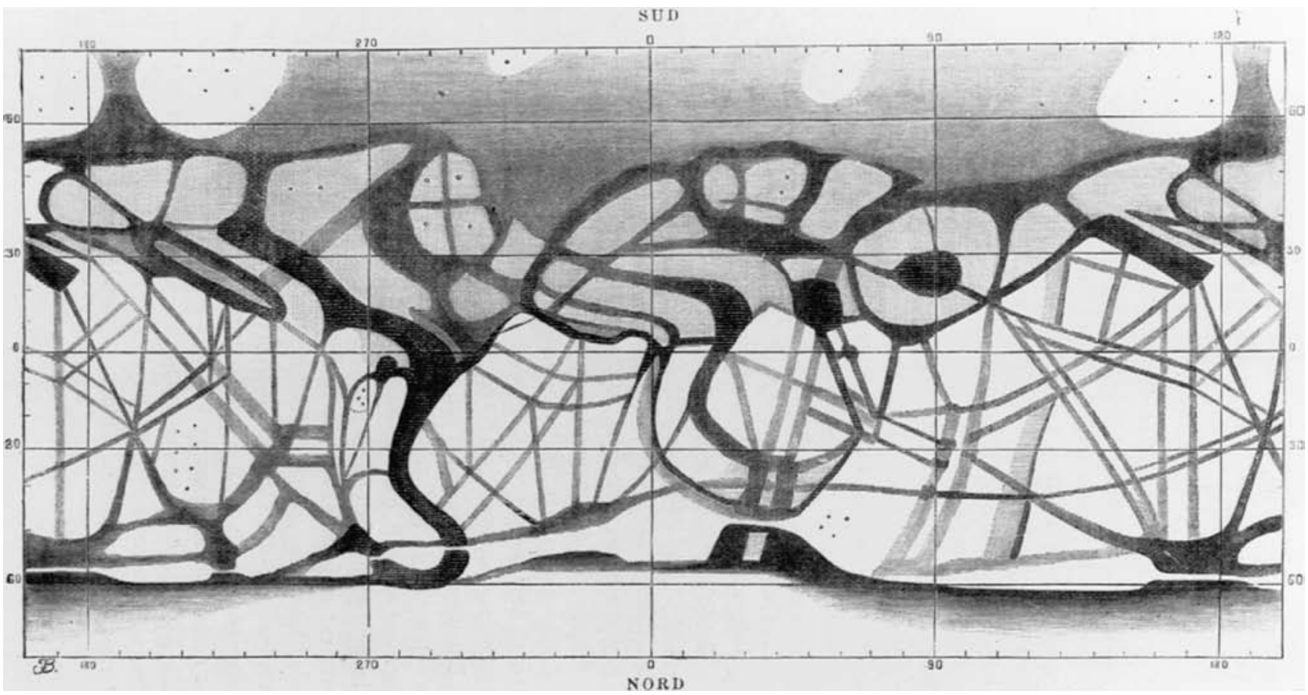


Fig. 8.5 Martian canals During the opposition of 1877, the Italian astronomer Giovanni Schiaparelli (1835–1910) mapped features he thought he saw on Mars, including a vast network of long, thin, straight lines criss-crossing the planet's surface. Some of the canals have apparently doubled, or divided in two, in this Mercator projection drawn by Schiaparelli during the opposition of 1881 using an 8.6-inch (22-centimeter) refractor. At this time, the apparent diameter of Mars was only 16 seconds of arc. Schiaparelli named these features *canali*, and they were likened to man-made water canals by subsequent observers, including Camille Flammarion (1842–1925) and Percival Lowell (1855–1916). Nevertheless, most astronomers failed to see the canals, and spacecraft have not detected them.

Table 8.3 Orbital missions to Mars[a]

Mission	Launch date	Orbit insertion	Discovery and/or accomplishments
Mariner 9	30 May 1971	13 Nov. 1971	Volcanoes, canyons, outflow channels
Viking 1	20 Aug. 1975	19 June 1976	Surface photographs, water ice in north polar cap
Viking 2	9 Sept. 1975	7 Aug. 1976	Same as *Viking 1*
Mars Global Surveyor	7 Nov. 1996	12 Sept. 1997	Laser altimeter, high-resolution images, gullies, ancient magnetism
2001 Mars Odyssey	7 Apr. 2001	24 Oct. 2001	Search for chemical, mineralogical and geological evidence for ancient water flow, vast amounts of subsurface water ice
Mars Reconnaisance Orbiter	12 Aug. 2005	1 July 2006	Water-related minerals, ancient water flow, gullies, high-resolution images

[a] Between 1971 and 1974 the Russians sent four *Mars* spacecraft into orbit around the red planet, with landers that crashed, missed the planet or failed within seconds of touchdown.

Table 8.4 Lander missions to Mars

Mission	Landing date	Site	Latitude (deg. N)	Longitude (deg. E)
Viking 1	20 July 1976	Chryse Planitia	22.27	311.81
Viking 2	3 Sept. 1976	Utopia Planitia	47.67	134.04
Mars Pathfinder	4 July 1997	Ares Vallis	19.09	326.51
MER Spirit[a]	4 Jan. 2004	Gusev crater	−14.57	175.47
MER Opportunity[a]	25 Jan. 2004	Meridiani Planum	−1.95	354.47
Phoenix	25 May 2008	North polar plains	68.22	234.25

[a] The acronym *MER* denotes the *Mars Exploration Rover*.

Table 8.5 Composition of the atmosphere at the surface of Mars[a]

Species	Abundance	Species	Abundance
Carbon dioxide (CO_2)	95.32 percent	Water vapor (H_2O)	0.016 percent[b]
Nitrogen (N_2)	2.7 percent	Neon (Ne)	2.5 ppm
Argon (Ar)	1.6 percent	Krypton (Kr)	0.3 ppm
Oxygen (O_2)	0.13 percent	Xeon (Xe)	0.08 ppm
Carbon monoxide (CO)	0.07 percent	Ozone	(0.04 to 0.2)[b] ppm

[a] Composition by volume in percent or in parts per million, denoted ppm. Because carbon dioxide varies seasonally due to condensation at the polar caps, all percentage abundances will vary seasonally. Adapted from H. H. Kieffer, B. M. Jakosky, C. W. Snyder, and M. S. Matthews (eds.), *Mars*, University of Arizona Press, Tucson, 1992.

[b] The abundance of water vapor and ozone vary with season and location. The annual global average of water vapor is 0.016 percent by volume.

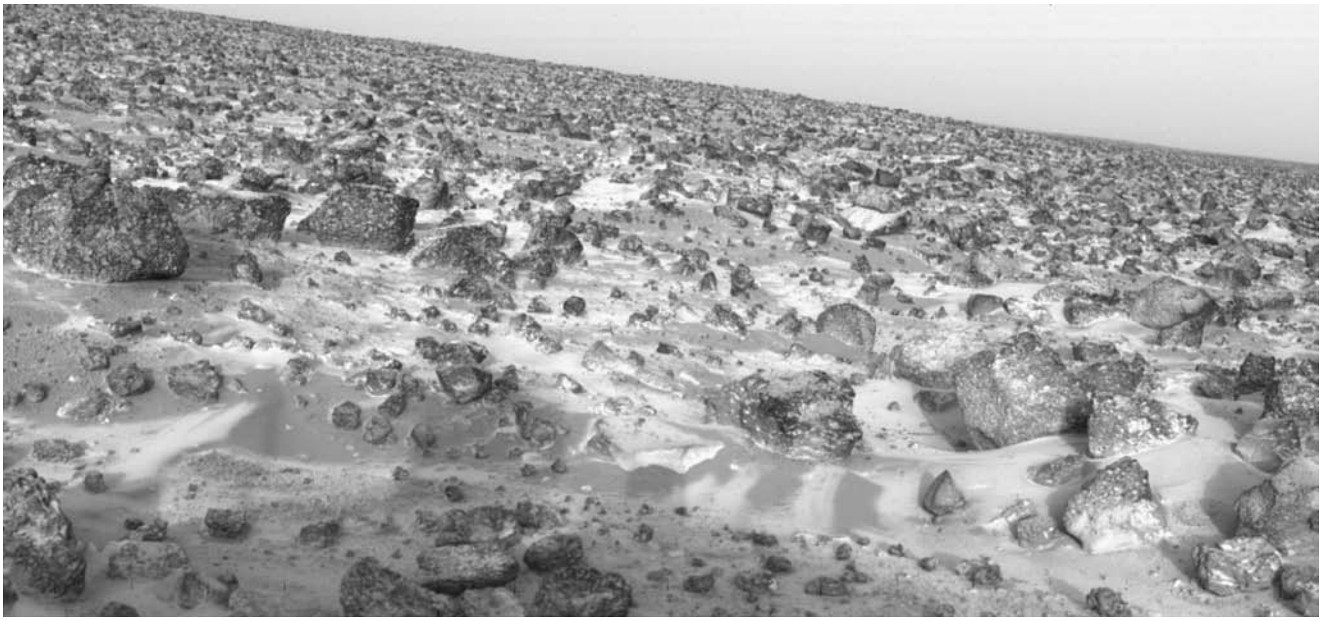


Fig. 8.6 Frost on Mars Atmospheric water vapor freezes onto the surface of Mars, producing a very thin coating of water ice on rocks and soil photographed from the *Viking 2* lander at its Utopia Planitia landing site on 18 May 1979. Scientists believe dust particles in the atmosphere pick up bits of solid water; carbon dioxide, which makes up 95 percent of the Martian atmosphere, freezes and adheres to the particles and they become heavy enough to sink. Warmed by the Sun, the surface evaporates the carbon dioxide and returns it to the atmosphere, leaving behind the water and dust in the white patches of frost shown here. The frost remained on the surface for about 100 Earth days. (Courtesy of NASA/JPL.)

argon (1.6 percent), while there is almost no oxygen (0.13 percent). In contrast the Earth's atmosphere is 77 percent nitrogen and 21 percent breathable oxygen.

The small amount of oxygen that is now present in the Martian atmosphere is the by-product of the destruction of carbon dioxide by energetic sunlight. This process also results in the production of exceedingly small amounts of ozone. Since there is so little ozone in the Martian atmosphere, it has no ozone layer, and the planet's surface is exposed to the full intensity of the Sun's lethal ultraviolet radiation. By way of comparison, the Earth has a thick ozone layer high in its atmosphere, which absorbs most of the dangerous ultraviolet sunlight and keeps it from reaching the ground.

There is now very little water vapor in the Martian atmosphere, about 0.016 percent near the surface. And that is about as much of the vapor that the tenuous atmosphere can hold. It is practically saturated with water vapor. When the temperature drops, water can condense and freeze out of the saturated air, forming low-lying mists or ground fogs in canyons and frosts on the surface (Fig. 8.6).

The concentrations of Martian water vapor vary with location and season, but they are always low. If all the water vapor above a given place on Mars could rain down to the surface as a liquid, the average depth would be 0.000 015 (1.5×10^{-5}) meters. By comparison, the Earth's atmosphere normally contains 10 000 times as much water vapor, capable of raining onto the ground with depths of several centimeters. Moreover, the pressures and temperatures on Mars are such that it cannot now rain on Mars.

The thin, cold Martian atmosphere

Since Mars has only a tenth of the Earth's mass, the red planet has less gravitational pull. It might thus be expected to retain a less substantial atmosphere. And because Mars is 50 percent farther from the Sun than Earth, it receives about half as much sunlight and ought to be colder. The ground-level pressure and temperature on Mars are, in fact, lower than those outside the highest-flying jet airplane on Earth. The atmosphere near the surface of Mars is almost as thin as our best laboratory vacuums, and the temperatures are usually below the freezing point of water.

The surface pressure at the bottom of the thin Martian atmosphere was first accurately determined when *Mariner 4* passed behind the planet, and its radio signal penetrated the Martian atmosphere in order to reach Earth. The measurements indicated that the atmosphere at the surface of Mars is about 150 times thinner than the one on Earth, or that the average Martian surface pressure is about 0.007 bar or 7 millibar, at about one hundred-fiftieth that of Earth's atmosphere at sea level, pegged at 1.000 bars. As on the

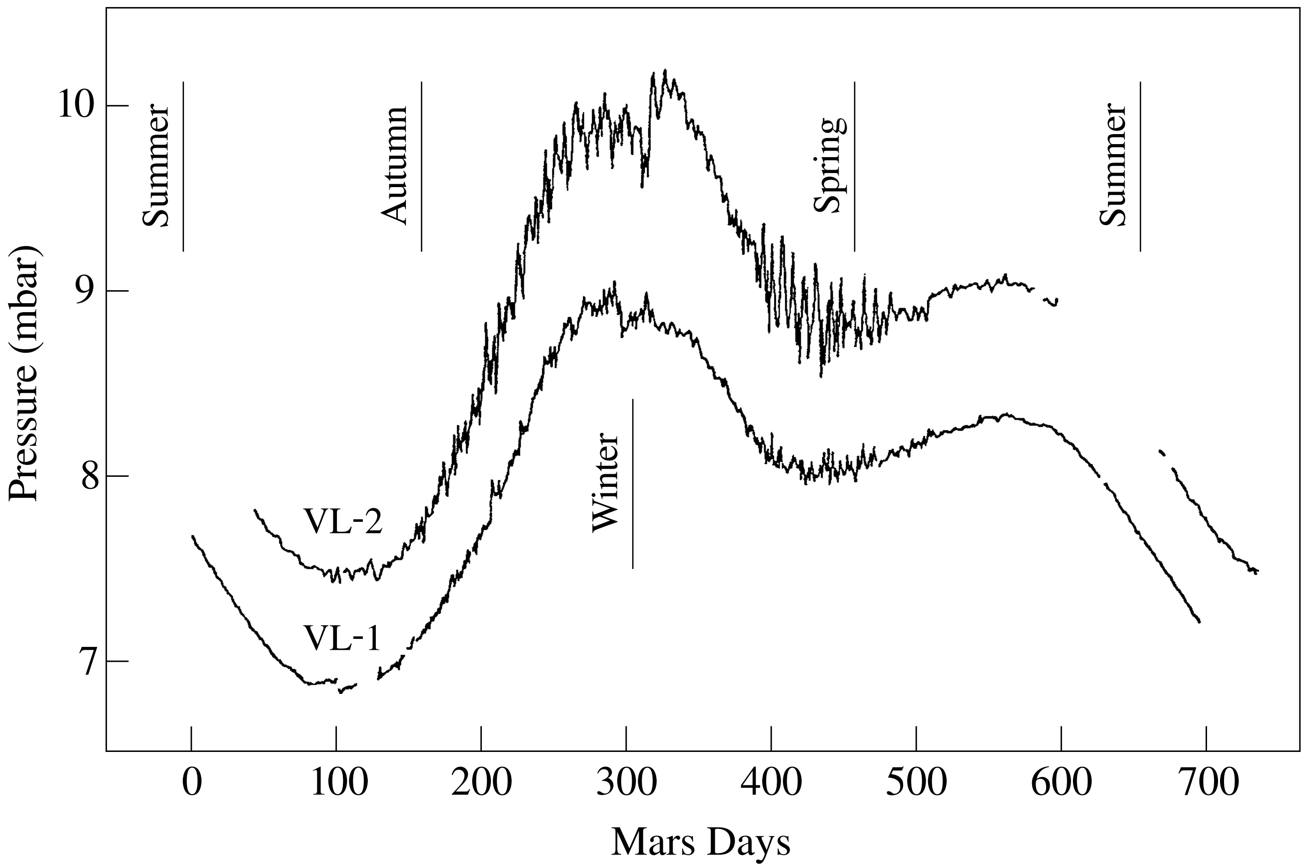


Fig. 8.7 Martian surface pressure Daily mean surface pressures at the two *Viking Lander* (VL) sites recorded for one Martian year, showing that the red planet periodically removes and replaces large amounts of carbon dioxide in its atmosphere. The seasons are for the northern hemisphere, and the pressure is given in millibars (mb), where 1 mb = 0.001 bar and 1 bar is the sea-level pressure of the Earth's atmosphere. At the *Viking 1* site (*bottom curve*), the pressure ranged from 6.7 mb during the northern summer to 8.89 mb at the commencement of northern winter. At the *Viking 2* site (*top curve*) the equivalent data were 7.4 mb and 10 mb. The higher values are probably due to the lower elevation; there was an approximate 1.1-kilometer difference in the elevation of the two landing sites. Dust storms are thought to produce some of the smaller non-seasonal pressure variations. The seasonal pressure differences seem to be dominated by the southern polar cap, which is larger than the northern one. In the southern winter, and northern summer, carbon dioxide is condensed out of the atmosphere, enlarging the southern polar cap and reducing the total surface pressure of the planet. In southern summer, and northern winter, the carbon dioxide has been released back into the atmosphere, with an increase in the total surface pressure.

Earth, the atmospheric pressure on Mars depends on the height of the surface, ranging from about 14 millibar in the bottom of the deep Hellas basin to about 3 millibar at the top of the tallest volcano.

The Martian surface pressure also varies with the season, as the polar caps grow or recede. The two *Viking* landers found, for example, that the surface pressure on Mars can change seasonally by 30 percent, while always remaining less than one-hundredth of that on Earth (Fig. 8.7). The frozen world becomes so cold during the southern winter that almost a third of its atmosphere freezes, dropping out of the sky and enlarging the southern polar cap. When the temperature rises in southern summer, the gas is released back into the atmosphere and the surface pressure rises again. This enormous seasonal change in the mass of the atmosphere has no counterpart on Earth, where pressure changes of a couple of percentage points are cause for concern, often indicating a major storm.

Without a thick atmosphere to trap solar heat, there is no pronounced greenhouse effect on Mars and temperatures at the planet's surface are only a few kelvin warmer than that expected from direct sunlight. The surface temperatures vary from day to night, from summer to winter, and by proximity to the equator or poles. Even in the summer at latitudes below 60 degrees north or south, typical surface temperatures range from 180 kelvin at night to 290 kelvin at midday. Even though temperatures of

Fig. 8.8 Dunes in Endurance crater on Mars Dust accumulates on the crests of dunes found on the floor of Endurance crater. This false-color image, taken from the *Mars Exploration Rover*, *Opportunity*, exhibits a "blue" tint in the flat surfaces between the dune flanks. This tint may be due to mineral-containing, sphere-like "blueberries," which accumulate on the flat surface and may be formed by ancient flows of liquid water. (Courtesy NASA/JPL/Cornell U.)

the immediate surface rise above the freezing point of water of 273 kelvin at low latitudes near midday, above freezing temperatures occur only within a thin upper crust; at a depth of just a few centimeters below the surface the mean temperature remains at 210 to 220 kelvin. At the poles in winter, temperatures drop to 150 kelvin, permitting carbon dioxide to condense out of the atmosphere.

Because the surface temperature and pressure are so low, liquid water cannot now stay on the surface of Mars. Over most of the surface, the temperature is usually below the freezing point of water, and when it warms above freezing, in the midday summer sunlight and near the equator, the water turns almost directly into vapor.

Water will stay liquid only if it is warm enough and at high enough pressure. Drop the temperature, and it will freeze; drop the pressure, and it will turn straight into water vapor, a process called *sublimation*. At the relatively low pressure that exists over large regions on Mars, any exposed liquid water would vaporize away. If liquid water were released onto the surface from a deep, possibly warmer interior, that water would survive for just a brief time before freezing into ice or evaporating and turning into water vapor.

The winds of Mars

Mars experiences substantial winds driven by temperature differences in the thin atmosphere. As on Earth, the Martian atmosphere is warmed by sunlight during the day and cools at night, but due to the absence of a thick moderating atmosphere or oceans, these daily temperature variations are extreme on Mars. Since the Sun also warms the summer hemisphere more than the winter one, Mars has global temperature differences that depend on the season, just as the Earth does.

The Martian atmosphere responds to both the daily and seasonal temperature differences by generating winds that blow from hot to cold regions, transporting heat and trying to equalize the temperatures. And since a rise in temperature is equivalent to an increase in pressure, it is high-pressure air that rushes toward low pressure, creating the wind.

The influence of the winds on Mars is pervasive. They stir up dense, billowing clouds of dust and scour the surface, creating time-varying light and dark patterns and carrying dust from one place to another. These are called *aeolian effects*, after Aeolus, the Greek god of the wind.

The tiny dust particles may be hazardous to future visitors to Mars. The dust will gum up spacesuits, scratch helmet visors, cause electrical shorts, sandblast instruments and clog motors. On the Moon, which is similarly dusty but has no substantive atmosphere or winds, spacesuits lasted only two days before they began to leak, and unlike the Moon, intense dust storms can arise on Mars, creating blizzards that will limit vision and drive dust into an astronaut's clothing and equipment.

The icy Martian winds have swept up vast dunes of sand and fine-grained dust. Rippled dunes have piled up in basins, chasms, and craters (Fig. 8.8), but the dunes cover only a small percentage of the land on Mars, probably less than 1 percent. Starkly beautiful patterns are created when the polar caps warm up in local spring and summer, exposing dark sand dunes.

As daily temperatures rise, winds stir up small, local dust storms, in much the same way that winds sometimes whip the terrestrial soils into towering columns called dust

Fig. 8.9 Twisted paths of dust devils on Mars Spinning columns of warm air, called dust devils, rise above the Sun-heated surface of Mars. Each tornado-like vortex picks up light-colored dust, exposing the darker surface underneath. Dust devils have created this wild pattern of criss-crossing dark streaks in the rippled flats of Argyre Planitia, covering an area 3 by 5 kilometers at latitude of 51 degrees south. This image was taken from the *Mars Global Surveyor* in March 2000. (Courtesy of NASA/JPL/MSSS.)

devils. These local dust storms form when the ground heats up during the day, warming the air immediately above the surface. The warm air rises in a spinning column that moves across the landscape like a miniature tornado, sweeping up dust that makes the vortex visible and leaving a dark streak behind. They have scratched tangled paths across some parts of Mars, often crossing hills and running across large sand dunes and through fields of house-sized boulders (Fig. 8.9).

Small dust storms can form simultaneously at several points on Mars, and then coalesce with each other, veiling the entire planet in a vast planet-wide dust storm, far larger than any seen on Earth. These awesome, globe-encircling storms occur during the hot summers in the southern hemisphere when Mars comes closest to the Sun and the surface temperatures are the highest. The rapid temperature increase generates hurricane-speed winds that can sweep fine dust particles high into the atmosphere. As more dust is carried aloft, it absorbs sunlight, further heating the atmosphere, strengthening the winds, and eventually covering the planet with a deep opaque cloud of dust (Fig. 8.10). But as the darkening cloud blots out the Sun, the lower layers of the atmosphere cool, the winds diminish, and the dust settles back down to the surface.

Such global dust storms are not generated during cool, long summers in the northern hemisphere, when Mars is 20 percent further away from the Sun than in southern summer, receiving less heat while also moving at a slower speed. It is the extra sunlight during the southern summer, when Mars is close to the Sun, that provides the heat and winds that energize the most powerful storms.

8.5 The polar regions of Mars

Seasonal polar caps on Mars

As on Earth, there are large white caps at both poles of Mars, first observed with telescopes centuries ago. These seasonal deposits change in size during the Martian year, growing in the cold of local fall and winter and shrinking in the summer heat. But these varying caps are poles apart in size and seasonal change. The polar cap in the southern hemisphere is larger in the local winter than the seasonal northern cap ever gets, and the southern cap is smaller in the local summer than is the northern one.

This asymmetry is a direct consequence of the planet's eccentric orbit, which carries Mars closest to the Sun, at perihelion, in the southern summer and farthest away during the southern winter, which is longer and colder than the northern one. The relative warmth of southern summer near perihelion causes the south polar cap to almost disappear from sight while the relative cold of southern winter produces a larger polar cap. The north polar cap grows to a lesser extent in the shorter, colder northern winter, when the red planet is closest to the Sun, near perihelion, and moving at its fastest orbital speed.

During the fall and winter, when the temperature drops below 150 kelvin at either pole, carbon dioxide condenses to form a seasonal cap composed of frozen carbon dioxide, or dry ice. This is the same dry ice that is used on Earth to keep ice-cream, lobsters and other things cold for days at a time.

In the spring and summer, when the polar temperature rises above 150 kelvin, the carbon dioxide cap sublimates, or evaporates, from solid ice to gas, returning to the atmosphere. This condensation of carbon dioxide during winter and its subsequent sublimation in the spring is what gives rise to the familiar waxing and waning of the Martian polar caps. The process is entirely analogous to the snowfall that blankets the Earth's polar regions in the winter and evaporates in the summer, except the "snowfall" on Mars mainly consists of dry ice with just a few snowflakes of water ice. It also accounts for the enormous seasonal change in the surface pressure on Mars.

The residual, remnant, perennial or permanent caps on Mars

At both poles, the caps never completely disappear in the heat of the summer, when the temporary, seasonal

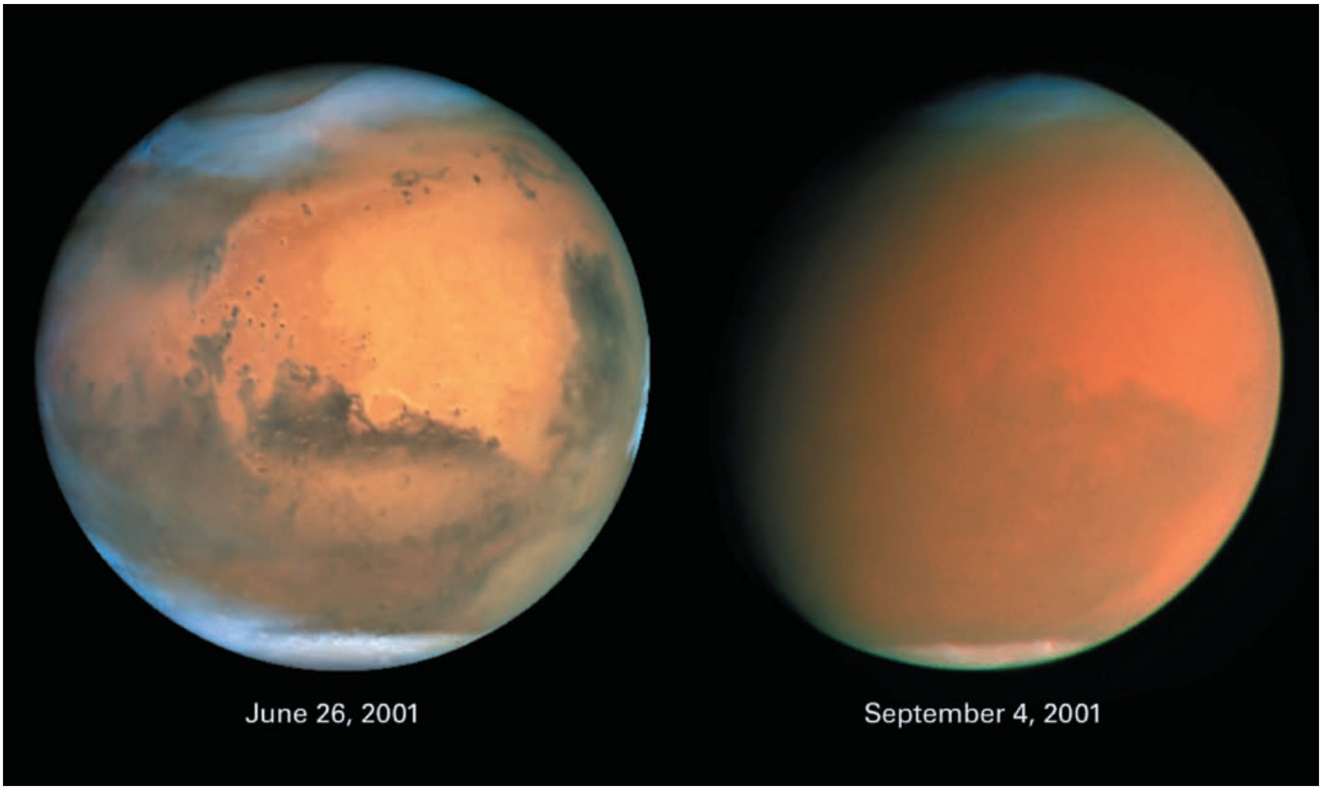


Fig. 8.10 Dust storm clouds out Mars Nothing on our world matches the global dust storms on Mars, dramatically displayed in this pair of natural-color *Hubble Space Telescope* images. Surface features that were crisp and clear when the first picture was taken (*left*) were covered with blinding dust by the time of the second picture (*right*). (Courtesy of NASA/JPL/STScI/James Bell/Michael Wolf/and the Hubble Heritage Team.)

deposits of dry ice sublimate back into the atmosphere. Residual, or remnant, polar caps are left behind. Since they remain throughout the Martian year, these residual caps have also been called perennial or permanent caps.

In the northern spring and summer, the central portion of the seasonal carbon dioxide cap sublimates completely to expose a water-ice cap (Fig. 8.11). As the temperature rises, water vapor is released above the north pole of Mars, which was how the *Viking* spacecraft showed that the north polar cap contains residual water ice underneath its seasonal dry ice covering.

In the south, the carbon dioxide cap does not dissipate completely and never entirely disappears, but high-resolution thermal and radar images of the south polar cap, from the *Mars Global Surveyor, 2001 Mars Odyssey* and *Mars Express* spacecraft, indicate that water ice lies beneath the remaining dry ice. The south polar region apparently contains enough frozen water to cover the whole planet in a liquid layer approximately 11 meters deep – if it ever melted. As in the north, the seasonal carbon dioxide cap in the south freezes into a winter layer about one meter thick on top of permanent cap of water ice.

At both poles, layered deposits several kilometers thick extend out to roughly 80 degrees latitude (Fig. 8.12). Individual layers are best seen in the walls of valleys cut into the sediments, where up to 20 layers have been exposed, each a few tens of meters thick and with strata as thin as 10 centimeters. The extensive regular polar layers are interpreted as dark mixtures of ice and dust separated by bright layers of nearly pure ice. They trace the history of periodic climate change at least for the past few million years and perhaps longer.

The repetitive polar layering is attributed to cyclical changes in the planet's tilt and orbital orientation, similar to the Milankovitch cycles that regulate the distribution of solar radiation on Earth, producing its ice ages. These astronomical rhythms are much larger in amplitude on Mars than on Earth, and are believed responsible for layers deposited with different dust-to-ice ratios in both the northern and southern polar caps. Radar reflections from different north polar ice layers, observed from the *Mars Reconnaissance Orbiter*, are consistent with theoretical models of how changes in the tilt of Mars' rotational axis have produced changes in the planet's climate over the past 4 million years continuing to the present.

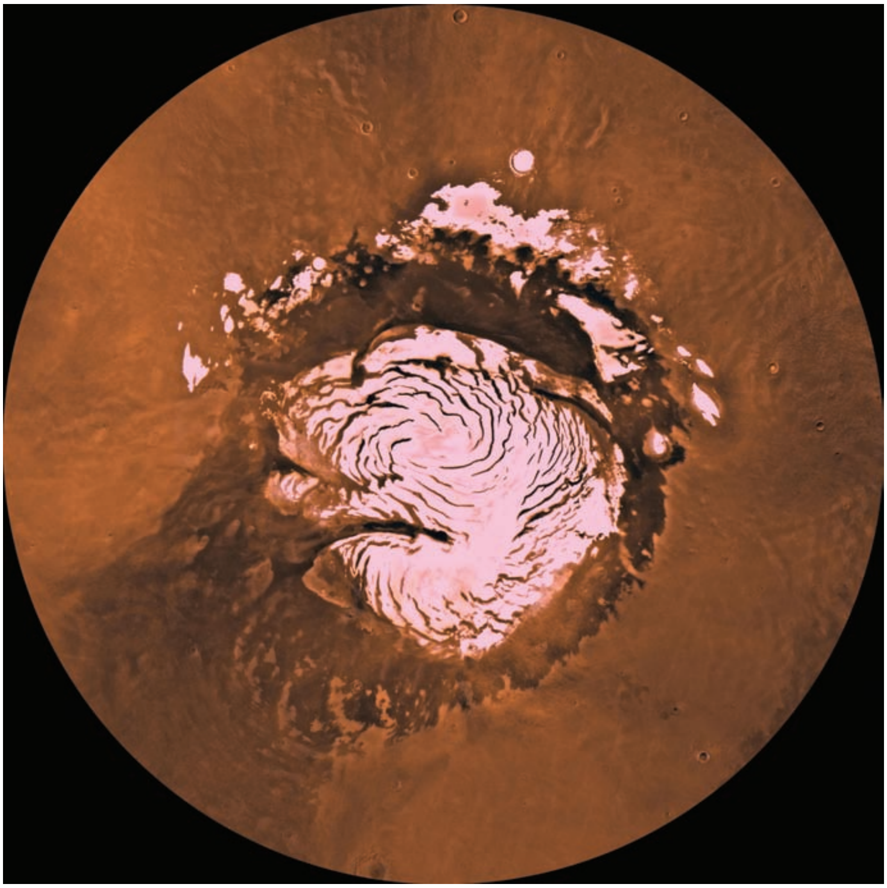


Fig. 8.11 Permanent, residual north polar cap of Mars The portion of the north polar cap that remains in summer is an imposing mountain of water ice, about 1200 kilometers across. The summit, which nearly corresponds with the planet's spin axis, stands about 3 kilometers above the flat surrounding plains. The north residual cap is surrounded by a nearly circular band of dark sand dunes formed and shaped by wind. This image, which was acquired by a *Viking* orbiter during the northern summer of 1994, strongly resembles that taken by the *Mars Global Surveyor* in the northern summer of 1999. Both the north and south residual caps contain deep valleys that curl outward in a swirled pattern that has been cut and eroded into the icy deposits, like a giant pinwheel. (Courtesy of NASA/JPL/USGS.)

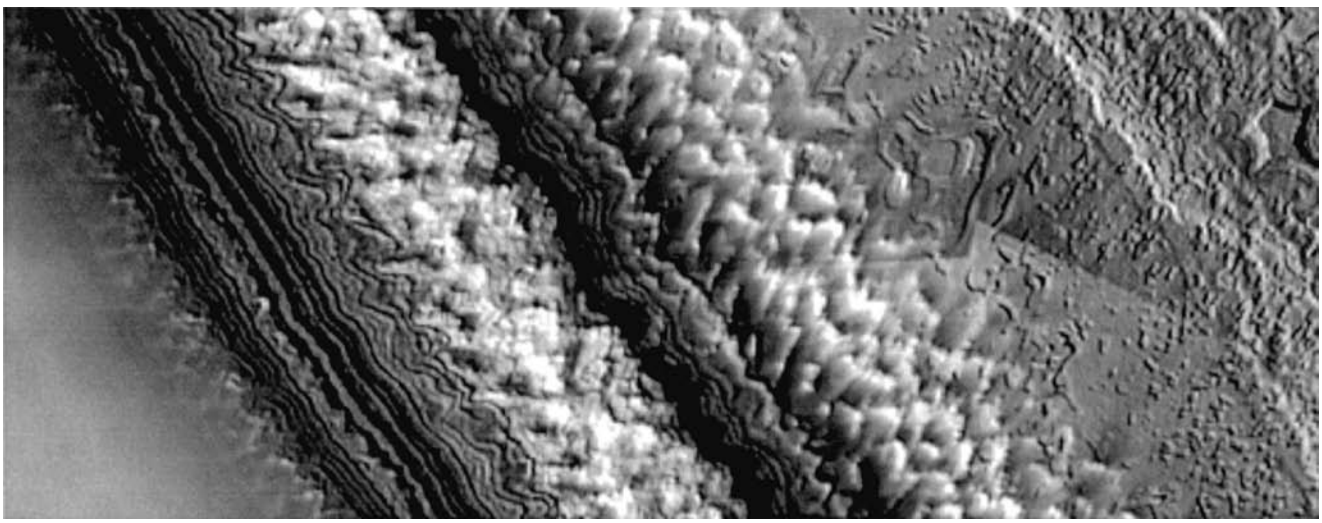


Fig. 8.12 Layered polar terrain on Mars These layers, exposed in the south polar residual cap of Mars, consist of bright ice and dark fine dust deposited over millions of years. The layered terrain in both the north and south residual caps is thought to contain detailed records of the climate history of Mars. This image covers an area of 10 by 4 kilometers. It was taken from the *Mars Global Surveyor* in October 1999 at 87 degrees south and 10 degrees west, near the central region of the residual south polar cap. (Courtesy of NASA/JPL/MSSS.)

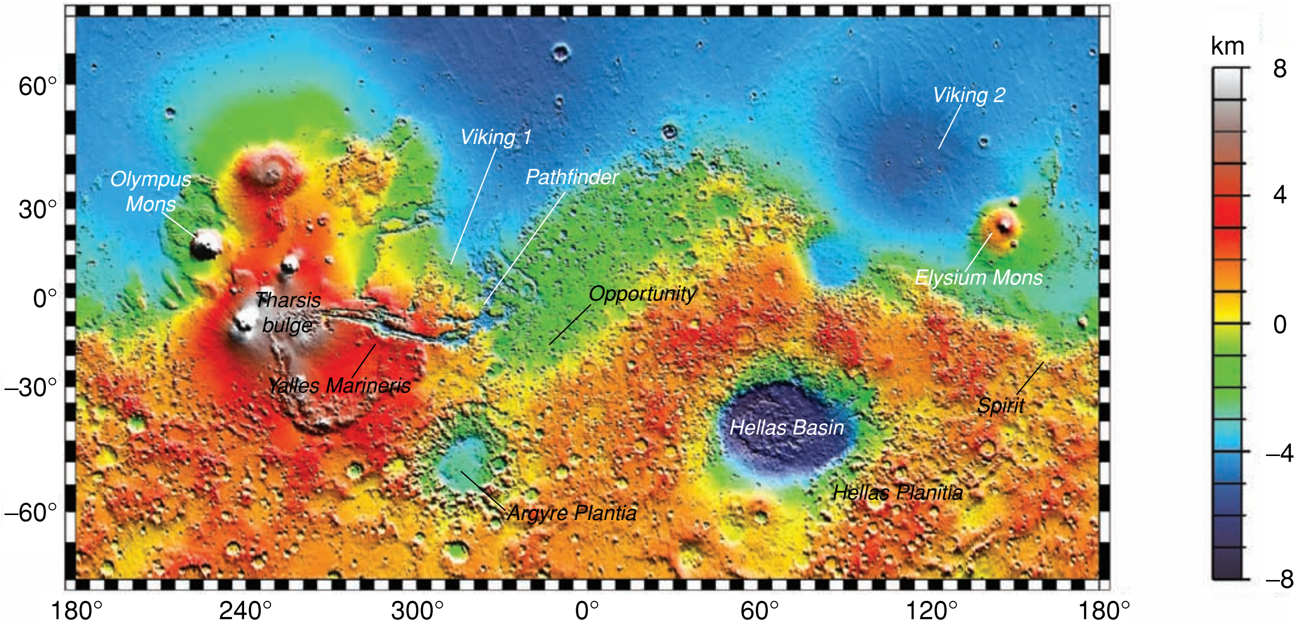


Fig. 8.13 Global topography of Mars By measuring the round-trip time of laser pulses bounced off the surface of Mars, the laser altimeter aboard the *Mars Global Surveyor* orbiter measured the altitude with a vertical accuracy of about 1 meter, providing this topographical map of Mars between latitudes of 65 degrees south (*bottom*) and 65 degrees north (*top*). The locations of five landing sites are labeled with the names of the landing spacecraft. This map portrays the great elevation difference between the northern, low-lying plains and the cratered southern highlands, and records the downhill direction that liquid water would flow. The red areas in the southern hemisphere are high regions, about 4 kilometers above the average surface height, and the blue regions of the northern hemisphere are low places, about 4 kilometers below the average height. The North Polar Basin, or Borealis Basin, is the large blue, low-lying area at the top of this topographical map; it covers about 40 percent of the planet. The map also shows the Tharsis bulge that lies near the Martian equator in the east longitude from 220 to 300 degrees. The bulge includes several major shield volcanoes, such as Olympus Mons. The huge Valles Marineris canyon system extends to the west of Tharsis. The giant Hellas impact basin, at 45 degrees south and 70 degrees east, is about 3000 kilometers across and lies about 9 kilometers deep. (Courtesy NASA/JPL/GSFC.)

8.6 Highs and lows on Mars

The crustal dichotomy of Mars

Mars is a divided world with a lopsided form, referred to as a great crustal dichotomy. The different age, height and crater density of the bottom, southern hemisphere and the top, northern hemisphere distinguishes the division. The planet's southern half is ancient, generally elevated, and rough, with a highly cratered surface that resembles the highlands on the Moon. The northern hemisphere, by contrast, consists mainly of young, lower-lying, smooth and flat plains with relatively few craters, not unlike the lunar maria. This hemispheric division has been most clearly revealed by a laser altimeter aboard the orbiting *Mars Global Surveyor* spacecraft, which analyzes pulses of laser light to measure the light time delay and distance between the spacecraft and the planet's surface (Fig. 8.13).

The cratered scars of impacting meteorites can be found all over Mars, but the craters are more densely concentrated in the southern hemisphere (Fig. 8.14). Like the lunar highlands, the extensive craters on Mars date back to an intense bombardment by meteorites early in its history, estimated at 3.8 to 4.0 billion years ago and known as the Noachian era on Mars.

The largest meteorites have gouged huge impact basins out of the Martian surface, throwing up mountains along their rims. They retain their classical designations made more than a century ago at the time of early ground-based telescopes – such as Argyre for the "silver" island at the mouth of the Ganges River, and Hellas, the Greek word for Greece. The giant Hellas basin, some 2300 kilometers across, is covered with white frost in southern winter, forming a brilliant white disk seen from Earth. The large craters, which are smaller than the impact basins, are named after astronomers and scientists who have studied the planet.

Most of the north is depressed by a few kilometers below the mean level on Mars, while the majority of the south is elevated by a few kilometers. The average elevation of the south is 5.5 kilometers above that of the north. But

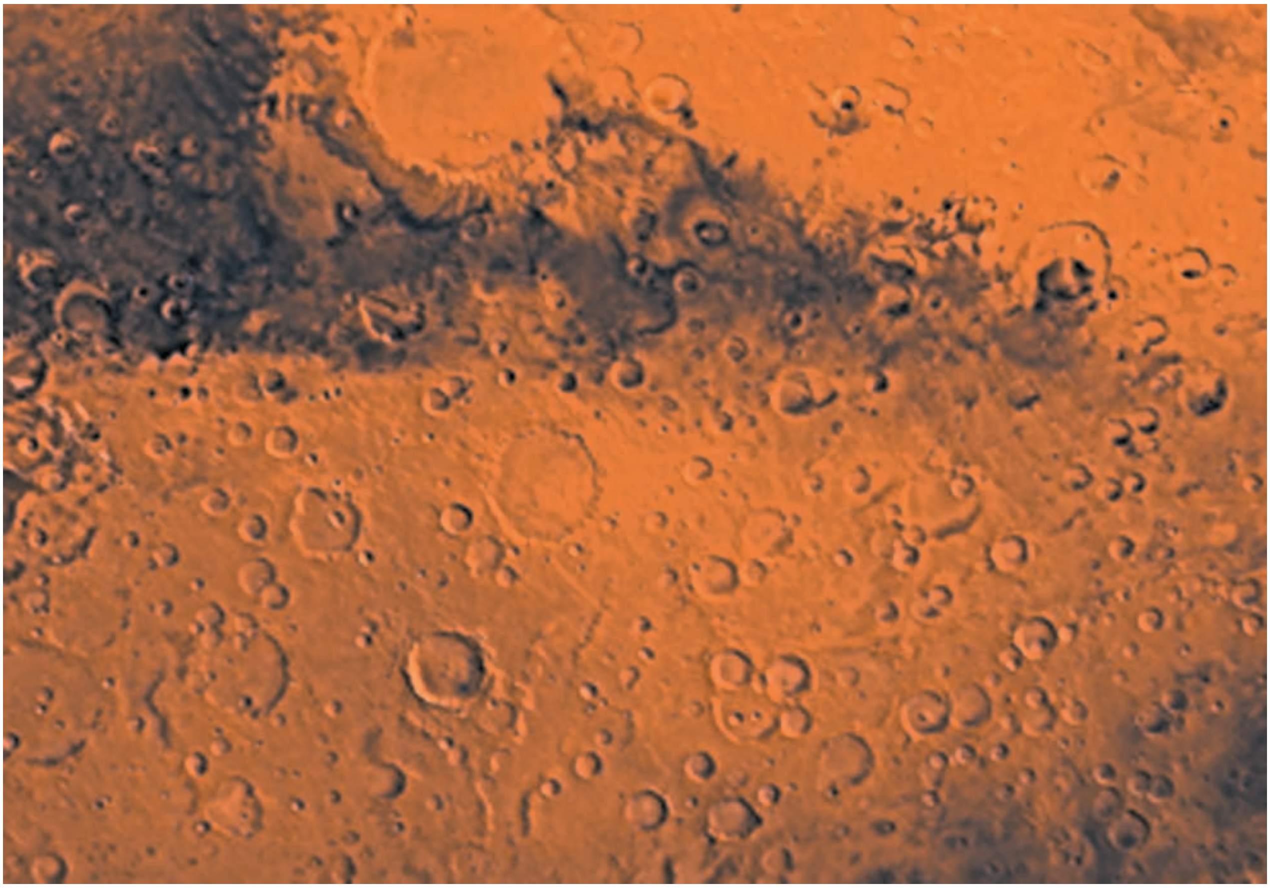


Fig. 8.14 Mars' Sinus Sabeus quadrangle Heavily cratered highlands dominate the Sinus Sabeus region of Mars, located just south of the equator between 0 and −30 degrees latitude. A large impact crater named after the Italian astronomer Giovanni Schiaparelli (1835–1910) marks the northern part of this mosaic image, taken from the a *Viking* orbiter. (Courtesy of NASA/JPL/USGS.)

there are exceptions, the north includes lofty volcanoes that rise as much as 27 kilometers above the main surface level of Mars and the south includes giant impact basins, such as the Hellas basin whose floor marks the lowest point on Mars. The northern half is also distinguished by the presence of volcanoes, canyons, flood channels, and extensive lava flows of volcanic origin.

The plains of Mars are designated by the names of lands, followed by the Latin *planitia*, meaning "a level surface or plain". But they are not completely smooth. Volcanoes rise up in some of them, mesas and buttes in others; boulders or dunes give them a small-scale texture.

The Latin term *planum*, meaning "plateau or high plain" designates flat elevated regions, in contrast to the low-lying planitia. Most of the planitiae are located in the northern hemisphere, while the plana are found just south of the equator (Fig. 8.15). Another Latin name, *terra*, is used to designate an extensive landmass in the older, heavily cratered highlands.

The plains of Mars are sparsely cratered, and therefore post-date the period of heavy bombardment that gave rise to the profusely cratered Martian highlands. Different plains nevertheless exhibit varying crater densities, and this provides a method for tracing the planet's development. Such comparisons show that different regions on Mars span a large range in ages, from the ancient, heavily cratered highlands to very lightly cratered volcanoes that may be younger than a million years.

The formation of cratered highlands, in the Noachian era between 4.0 and 3.8 billion years ago, was followed by the Hesperian era from 3.80 to 3.55 billion years ago, when vast canyons were formed and catastrophic floods carved out huge channels. The extensive lowland volcanic plains were emplaced across the Martian surface during the subsequent Amazonian era, after 3.55 billion years ago, leading to the formation of the Lunae Planum, Chryse Planitia, Syrtis Major Planum, Amazonis Planitia, and Utopia Planitia, with estimated ages of 3.5, 3.0, 2.9, 2.8 and 1.8 billion years, respectively. Tall volcanoes were also formed throughout this era, from 3.5 billion years ago to relatively recently, perhaps even now.

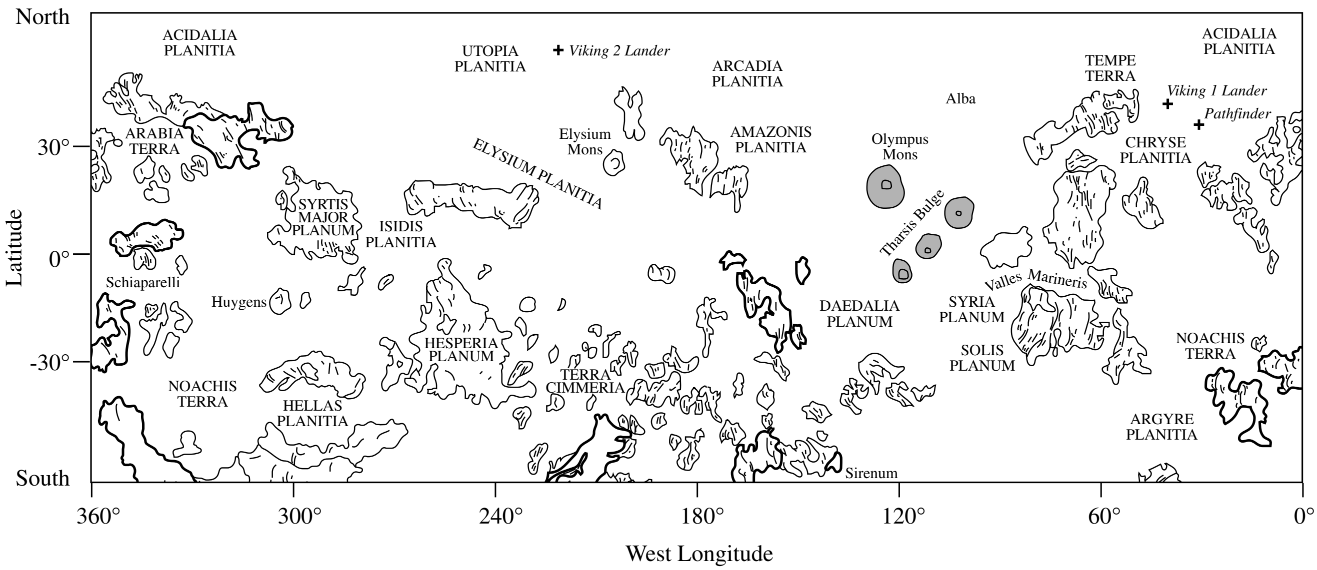


Fig. 8.15 Distribution of plains on Mars Low-lying volcanic plains, each designated as a planitia, are located throughout the northern hemisphere of Mars. Other relatively smooth regions are found at the top of elevated plateaus, each called a planum; they are located in the southern side of the equator. Small solid dots denote volcanoes, including Olympus Mons that rises out of the Tharsis bulge, or uplift. The small crosses designate the landing sites of *Viking 1* and *Pathfinder* in Chryse Planitia (*upper right*) and *Viking 2* in Utopia Planitia (*top center*).

The topographic mismatch, with highs on the bottom and lows on the top, probably arose soon after Mars formed, and well before volcanic material began to "resurface" the northern hemisphere. According to one theory, the dichotomy may have been created within, by some unevenness in the young planet's molten subsurface layers. Internal convection somehow became lopsided, creating a global imbalance in the sweep of material and the release of heat.

In a competing theory, an ancient impact created the crustal dichotomy from outside, soon after the planet formed and when collisions with large objects were still common. A Pluto-sized body then struck the northern hemisphere at a shallow angle, ripping off the crust. This theory has received support from instruments on the *Mars Reconnaissance Orbiter* and *Mars Global Surveyor*, which have provided detailed information about the elevations and gravity of the red planet. They have discovered a large basin in the northern hemisphere, named the North Polar Basin or the Borealis Basin, which covers 40 percent of the planet and is between 8500 and 10 600 kilometers across. Its low, flat and relatively crater-free topography, previously shown in Fig. 8.13, suggest that it was created by a single large impact. Similar impacts with relatively large bodies about 4 billion years ago may account for the formation of our Moon, by a glancing blow of a Mars-sized object with the young Earth, and the removal of the missing low-density crust of Mercury.

Soaring volcanoes and immense canyons on Mars

Powerful forces have molded the face of Mars at an unexpected scale, including huge volcanoes and immense canyon-lands that dwarf their terrestrial counterparts. The colossal Martian volcano Olympus Mons has an elevation of 27 kilometers above the mean surface level, about three times that of Earth's Mount Everest above sea level and 2.6 times the height of Hawaii's Mauna Kea above its base. A gigantic canyon, dubbed Valles Marineris after *Mariner 9*, could stretch from New York City to San Francisco, putting our Grand Canyon to shame. This imposing system of interconnected canyons, or chasmata, extends along the Martian equator for 4000 kilometers, one-fourth the way around the planet. In places the chasms are as wide and deep as Mount Everest is high.

Unlike the Grand Canyon, which formed over eons of erosion by flowing water, the Valles Marineris was created when Mars cracked open about 3.8 billion years ago. All the chasmata found in this region are huge surface cracks, unlike canyons found on Earth that have been formed by running water. After the formation of the Martian canyons, erosive forces took over. Gigantic landslides widened the canyon walls, howling winds roared through the canyons, and liquid water apparently once existed in parts of the abyss.

Superposed on the global dichotomy is the Tharsis bulge, which straddles the equator. It is more than

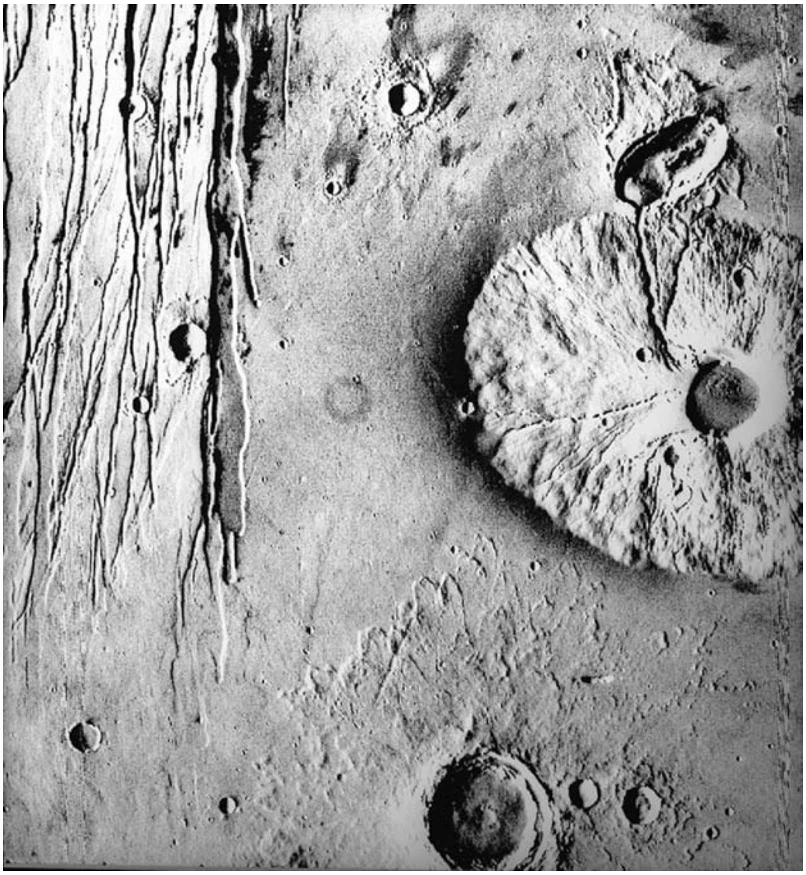


Fig. 8.16 Ancient Martian volcano Ceraunius Tholus The Tharsis region of Mars includes both volcanic and tectonic features. This *Viking 1* image portrays the shield volcano, Ceraunius Tholus (*center right*), which is 115 kilometers in diameter. A two-kilometer wide channel extends from the summit caldera down the flanks of the volcano through a crater and into the adjacent plains. Smaller channels are just visible elsewhere on the flanks. The surrounding region (*left*) includes intensely fractured terrain. These structures are probably related to the uplift of the Tharsis region, causing fracturing and faulting. The term *tholus* means "small domical mountain or hill," and *ceraunius* means "thunderclap." Ceraunius Tholus is named for the Ceraunii Mountains on the coast of Epirus, Greece. (Courtesy of NASA/JPL.)

5000 kilometers across and 10 kilometers high, overlying the ancient highlands and lowland plains near the equator. It was formed roughly 2 billion years ago, after the creation of the crustal dichotomy. The tallest volcanoes on Mars are located on the bulge, including the Olympus, Arsia, Ascreus and Pavonis Montes, which are all much larger than any terrestrial volcano.

With their gently sloping flanks and roughly circular summit calderas, the Martian volcanoes resemble the shield volcanoes of Hawaii, such as Mauna Loa, which has a similar slope but one-third the height and one-twentieth the volume of Olympus Mons. Such volcanoes are formed by the repeated eruption of lava that cascades down the flanks in thousand of individual flows.

Why are some Martian volcanoes so much higher than their terrestrial counterparts? Perhaps it has to do with the planet's outer shell, or lithosphere. Because Mars is smaller than the Earth, it probably cooled faster, and its lithosphere became relatively stronger and thicker, not breaking up into moving plates as in the Earth's lithosphere. This gives Martian volcanoes a longer chance to grow in one spot.

Since the lithosphere on Mars is one thick, solid plate, the planet's crustal movements are mainly vertical rather than horizontal. In contrast, the Earth's plates slide horizontally over the deep-seated hot-spot sources of magma. This motion limits the growth of individual shield volcanoes on Earth, and produces chains of smaller volcanoes, such as the Hawaiian chain in the Pacific Ocean. Large-scale plate tectonics almost started on Mars, but the development was stifled by the planet's rapidly cooling outer layers. The Martian crust therefore does not move across the internal hot spots, so the lava can erupt from them for billions of years, building up volcanoes far larger than any volcano on the Earth.

The difference in gravity between Mars and the Earth also contributes to the size of Martian volcanoes. As a volcano grows in height, it eventually becomes too heavy for the underlying rock to support, and the added weight causes the entire mountain to spread outwards. Because the force of gravity on Mars is only about one-third as great as that on Earth, the Martian volcanoes can grow more than twice as tall as their terrestrial counterparts before reaching the limiting height caused by too much weight.

The lack of craters on some Martian volcanoes, such as the summit of Arsia Mons and Elysium Mons, indicate that lava flowed from them as recently as 10 million to 100 million years ago. Yet Martian volcanism dates back billions of years. Some volcanoes erupted 2 or 3 billion years ago; Ceraunius Tholus is an example (Fig. 8.16). The

Table 8.6 Common features on the surface of Mars

Descriptor term	Feature type	Example
Chasma, Chasmata	Canyon, steep-walled trough	Candor Chasma (6°S, 71°W)
Labyrinthus, Labyrinthi	Complex of intersecting valleys or canyons	Noctis Labyrinthus (7°S, 101°W)
Mons, Montes	Mountain, volcano	Olympus Mons (18°N, 133°W)
Planitia, Planitiae	Low plain	Elysium Planitia (20°N, 230°W)
Planum, Plana	Plateau or high plain	Sinai Planum (15°S, 87°W)
Terra, Terrae	Extensive land mass	Noachis Terra (35°S, 335°W)
Vallis, Valles	Valley	Ares Vallis (10.4°N, 25°W)

Focus 8.1 Naming features on Mars

There are two names that describe every feature on Mars. They are a particular name that identifies it, plus a descriptive name that says what it looks like. Some of the descriptor terms, or feature types, commonly used on Mars are given in Table 8.6.

high crater density on the outer flanks and outer edges of the Martian colossus, Olympus Mons, indicates a similar old age for the edifice, even though the paucity of impact craters at its summit implies that lava flows may have occurred less than 300 million years ago. By way of comparison, the oldest volcanoes now on Earth have ages of just a few million years or less.

Since it is bigger than Mars, the Earth has remained much more internally active, continually renewing the terrestrial surface and destroying most of the Earth's older terrain. And the internal heat of the Moon and Mercury, which are smaller than Mars, cooled off long ago; they have not been active for billions of years. Mars is between these extremes. It is just large enough to have remained active for most of the solar system's history, but not so active that all record of its early history has been erased.

Names of the surface features on Mars

The system for naming all the surface features on Mars, which have largely been discovered after close-up scrutiny by instruments aboard orbiting spacecraft, is described in Focus 8.1.

The particular names for large craters on Mars are those of deceased astronomers or physicists, ranging from Tycho Brahe (1546–1601) to Gerard Kuiper (1905–1973), as well as writers who have contributed to the lore of Mars, such as Isaac Asimov (1920–1992), Edgar Rice Burroughs (1875–1950) and Robert A. Heinlein (1907–1988). Small craters receive the names of small towns of the world with a population less than 100 000. Large valleys, described by the Latin term *vallis*, are given the name of Mars in various languages. Examples include the Ares, Kasei, Nirgal, Simud and Tiu Valles, named respectively for the word "Mars" in Greek, Japanese, Babylonian, Sumerian and Old English. Small valleys receive the classical or modern names of terrestrial rivers.

Other features on Mars are designated for the nearest named albedo (light and dark) feature on the maps of Giovanni Schiaparelli (1835–1910) or Eugene Antoniadi (1870–1944), whose appellations were drawn from classical literature and the Bible. The bright areas were named for continents or islands, such as Argyre, Arabia, Chryse, Elysium, Hellas and Tharsis. The main dark areas were given the names of bodies of water, such as Sinus Sabaeus, Solis Lacus, and Syrtis Major. The complete list is available in the International Astronomical Union's *Gazetteer of Planetary Nomenclature*, located on the Internet at http://planetarynames.wr.usgs.gov/.

8.7 Flowing water on Mars long ago

Cold, parched and wrapped in a thin, carbon dioxide atmosphere, Mars today is a frozen, desiccated and inhospitable world. It cannot now rain on Mars, and liquid water cannot now remain on its surface. Yet several lines of evidence point to running water on Mars in the distant past, 3 or 4 billion years ago, when huge amounts of water swept through outflow channels and emptied into russet flood plains, creating streamlined, washed-out landforms, and stately rivers slowly carved deep, winding valleys into the surface. There may even have been ancient lakes or shallow seas on early Mars.

Orbiting spacecraft have recently detected surface minerals created by water flowing on Mars more than 3.5 billion years ago, and roving spacecraft have found additional evidence for the planet's watery past.

The very old, water-cut features that indicate liquid water once flowed across the Martian surface take two main

Fig. 8.17 Channels with tributaries on Mars Massive floods of water from the highlands into the Chryse basin in the lowlands may have carved these channels, located in the region of Mangala Vallis on Mars. The tributaries are rather shallow features, and join their main channels at quite acute angles. This image, taken from the *Viking 1* orbiter, has a width of 400 kilometers. (Courtesy of NASA.)

forms, known as the water networks and the outflow channels. The water networks, also known as valley networks and runoff channels, seem to have been derived from the gradual flow of liquid water. The immense outflow channels were gouged out of the surface by the powerful rush of short-duration floods.

Water networks on Mars

Networks of valleys cut across the oldest terrain on Mars, in the heavily cratered southern highlands, dating back to roughly 3.8 billion years ago. As the sinuous valley networks wind and meander downhill, they coalesce with several well-developed tributaries (Fig. 8.17), looking exactly like dry riverbeds on Earth. This suggests that the Martian water networks were not formed by rapid, surging floods.

The branching tributaries of some valley networks end abruptly in box canyons, suggesting that they formed by collapse into cavities formed by water running under the frozen, ice-rich surface. When viewed at high resolution, some individual valleys do not strongly suggest liquid flow, possibly due to subsequent modification of the surface. So some of the dry riverbeds, or water networks, might be best explained by collapse into features formed by underground rivers, and they may not require rain or running surface water. Instruments aboard the *Mars Reconnaissance Orbiter* and the *Mars Global Surveyor* have indeed revealed hundreds of small fractures in the equatorial regions of Mars, suggesting a network of ancient underground flows.

Ancient, water-charged torrents on Mars

Long, wide grooves have been gouged out of the equatorial regions of Mars, running downhill from the southern highlands into the northern lowland plains (Fig. 8.18). They tend to be narrow and deeply incised near their origins in the highlands and broad and shallow in the volcanic plains. Unlike the water networks, these enormous channels lack tributaries and are characterized by sculptured landforms such as scoured surface features, streamlined hills, and teardrop-shaped islands where the flowing water encountered an obstacle (Fig. 8.19).

Although some of the outflow channels date back almost to the end of the heavy bombardment 3.8 billion years ago, others were formed 2.5 to 3.5 billion years ago. That is relatively young on a cosmic scale, but still early in the planet's history and old in terms of geologic timescales.

The gigantic furrows bear all the marks of catastrophic outpourings of water, and are hence known as outflow

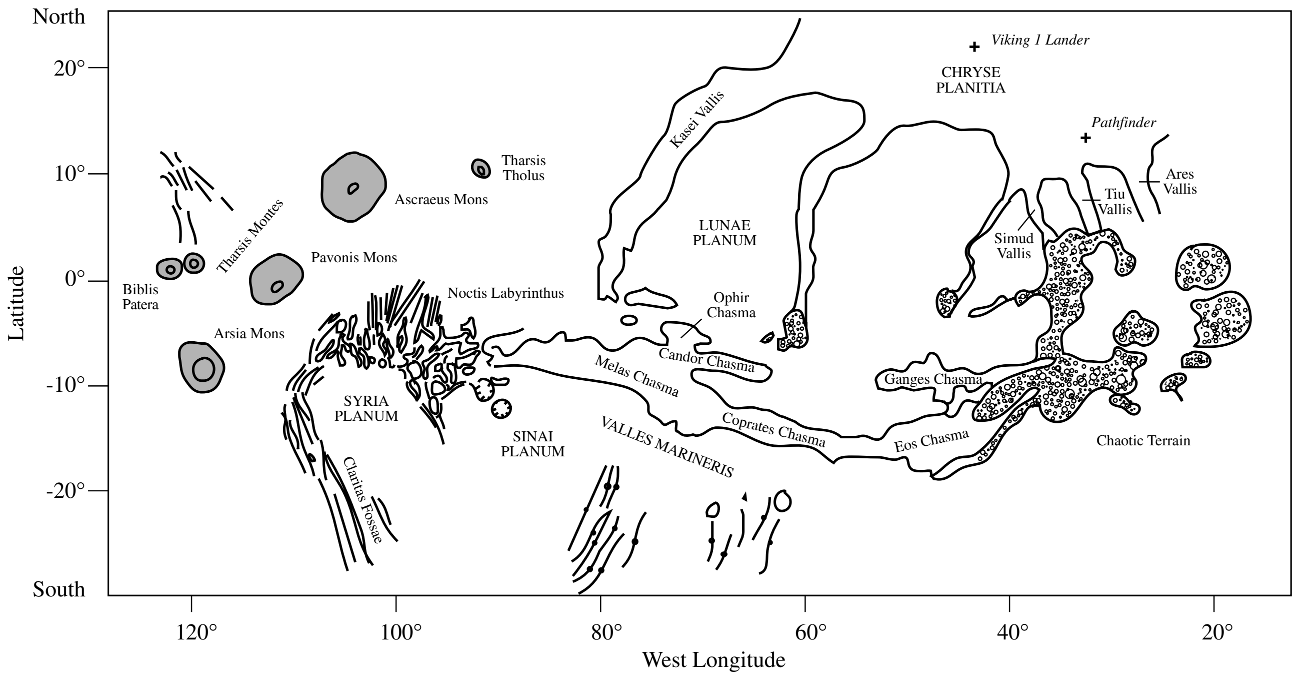


Fig. 8.18 Geological features of Mars The uplift that created the Tharsis Montes and nearby shield volcanoes (*left*) seems to have fractured the terrain and opened up an enormous network of chasmata, or canyons, knows as Valles Marineris (*center*). Catastrophic floods originating in the vicinity of these canyons flowed north (*top*) into Chryse Planitia, which contains the site of the *Viking 1* lander and the *Mars Pathfinder* lander.

Fig. 8.19 Streamlined island in outflow channel on Mars A raised crater rim acts as a barrier to the catastrophic floods that discharged from the outflow channel Ares Vallis. The water flowed from the southwest (*bottom left*) with a peak discharge more than 2000 times that of the Mississippi River. (Courtesy of NASA/Michael Carr.)

channels. They were formed by great, impulsive and short-lived floods of liquid water, somewhat like flash floods on Earth but on a more monumental scale.

A vast quantity of water was required to create the enormous outflow channels. They are sometimes more than 100 kilometers in width, up to 1000 kilometers in length, and as much as several kilometers deep. The discharge must have been enormous, flowing as rapidly as 75 meters per second and 1000 to 10 000 times faster than the Mississippi River on Earth. But we do not know how long the floods lasted, so we do not know the total volume of each flood.

Once released, the surging Martian torrents could not be stopped. Such discharges would not freeze even under present conditions on Mars. Large bodies of water must have been left in low-lying areas when the floods were over. Some scientists claim that Mars must have had oceans as extensive as those on Earth; others argue that seas larger than the Mediterranean were unlikely but that large lakes were possible. In either case, the lakes or seas would now be frozen.

Possible ancient lakes and seas on Mars

As we all know, water collects within holes in the ground, ranging in size from potholes in winter roads to stream-fed lakes and ocean basins. And if water once flowed across the surface of Mars, it would similarly pool in low-lying depressions, such as impact craters and basins, deep canyons and the northern lowland plains. At one time, they could all have been filled with water, forming ancient lakes and seas with perhaps a thin layer of ice on top, but all that now remains is their dried-out floors and sediment.

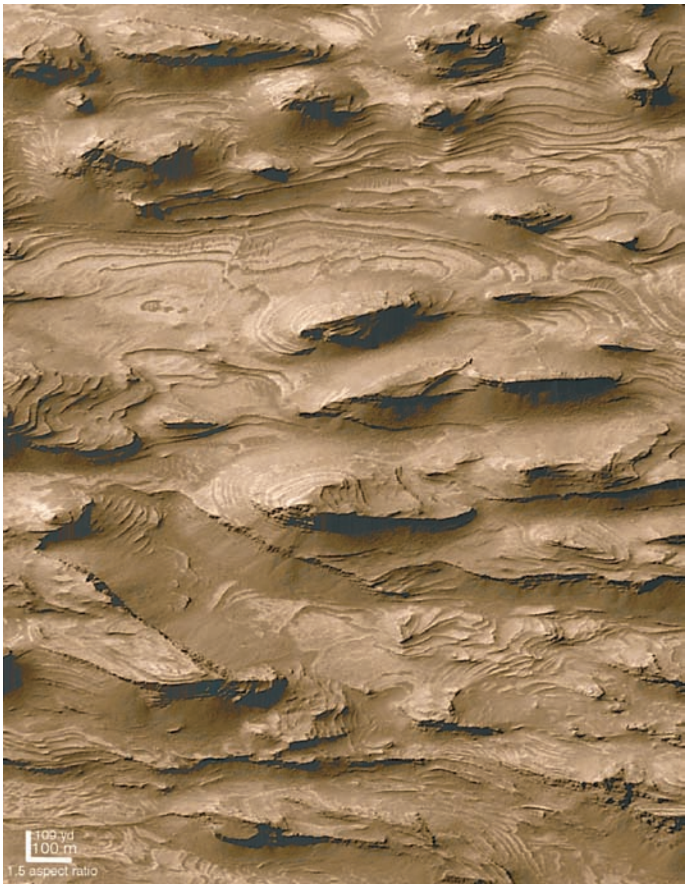


Fig. 8.20 Layered deposits in Mars' Candor Chasma The high-resolution camera aboard the *Mars Global Surveyor* orbiter revealed the presence of layered material in the floor of western Candor Chasma, at the far end of the main depressions within the extensive Valles Marineris. The numerous uniform deposits resemble regularly layered sedimentary rocks found on Earth. The Martian features could therefore be due to sediments that settled out of liquid water in ancient lakes or shallow seas a few billion years ago. Other layered deposits found in craters or basins may have existed before the canyons opened up in the surrounding terrain, and might be due to deposits of airborne dust settling out of the atmosphere, that were later buried and compacted, or to layers of volcanic material. (Courtesy NASA/JPL/MSSS.)

Widespread, stratified rock structures, found in topographical lows on Mars, could have been deposited by standing bodies of water. The layered material is located in impact craters, on parts of the Hellas impact basin, and on the floors of canyons in Valles Marineris, such as Candor Chasma (Fig. 8.20).

Thick sequences of hundreds of horizontal, regularly layered deposits are present in many places throughout the canyons, which seem to have contained lakes 3 to 4 billion years ago. The layered sediments were presumably deposited in these lakes, and then possibly compressed and cemented into rock. These lakes are supposed to have subsequently drained from the canyons into several large outflow channels, but the source of the sediment and the cause of the regular layering are still uncertain.

The low, flat northern regions of Mars could mark the dried-out bottom of a former ocean that once occupied up to one-third of the surface area of Mars, perhaps filling some or all of the North Polar Basin. Massive rivers that flowed through the outflow channels and into the northern plains might have fed the ocean. Possible ocean shorelines have been found, extending for thousands of kilometers and varying in elevation by several kilometers, perhaps due to the movement of Mars' rotation axis.

Orbiting spacecraft detect water-related minerals on Mars' surface

The infrared heat emission from the Martian surface has been imaged with spectrometers aboard orbiting spacecraft, revealing water-related deposits. The *Mars Global Surveyor* orbiter has, for example, mapped the presence of hematite, an iron oxide mineral that typically forms in the presence of water (Fig. 8.21). *Opportunity*, the *Mars Exploration Rover*, was subsequently directed to one of these locations of concentrated hematite, confirming the presence of flowing water in that location about 3.7 billion years ago.

The ancient southern highlands of Mars display a variety of mineral evidence for past water flow. An imaging spectrometer aboard the *2001 Mars Odyssey* orbiting spacecraft has found hundreds of deposits of chloride salts that typically lie in topographic depressions. They are most likely places where water was once abundant and then evaporated, leaving the salt deposits behind, perhaps 3.5 to 3.9 billion years ago.

More recently, the *Mars Reconnaissance Orbiter* has detected vast regions in the southern highlands that contain clay-like minerals, which can only form in the presence of liquid water. Volcanic lava buried the ancient regions

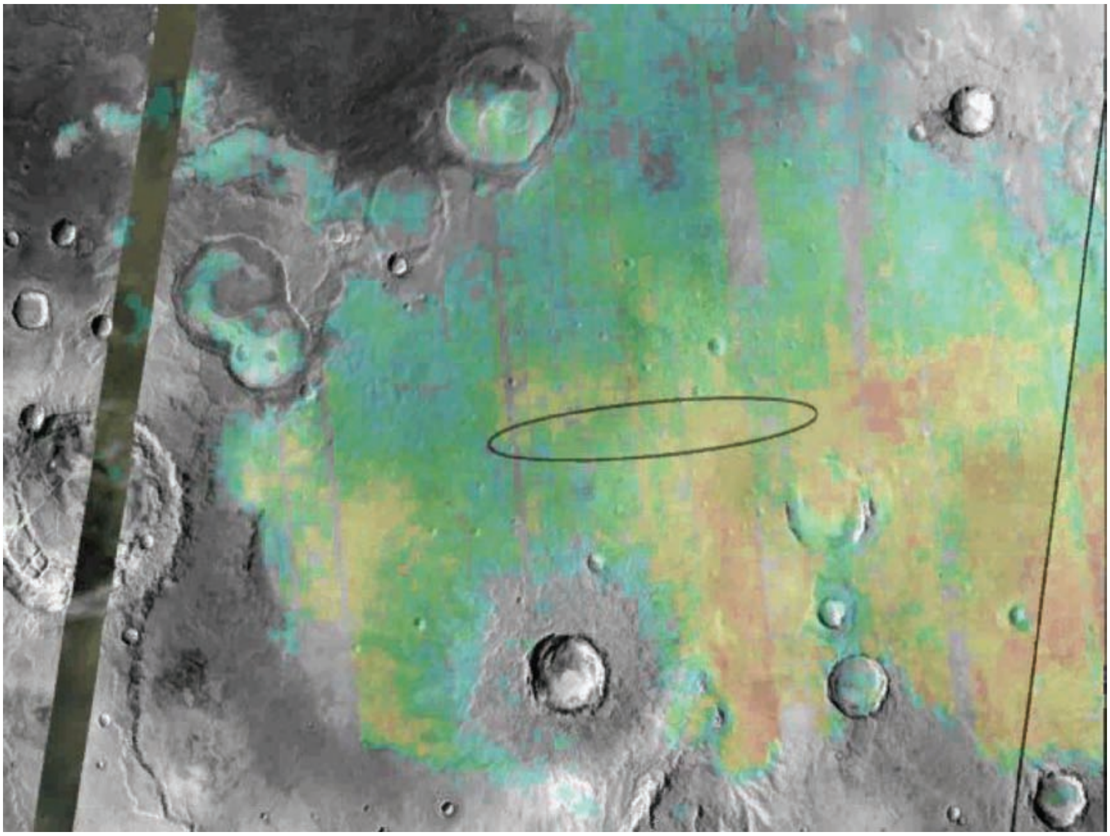


Fig. 8.21 Targeting water-containing minerals on Mars This image shows the abundances and locations of the mineral gray hematite at the landing site of the *Mars Exploration Rover*, *Opportunity*, in Meridiani Planum. It was targeted to land within the oval, which is about 71 kilometers long. A colored map from an instrument aboard the *Mars Global Surveyor* orbiter displays high (*red and yellow*) and low (*green and blue*) concentrations of hematite, an iron-oxide mineral that typically forms in the presence of liquid water. The underlying surface image that includes the adjacent craters is from the *2001 Mars Odyssey* orbiter. The *Opportunity* rover found abundant evidence near its landing site for flowing water in the ancient past, about 3.7 billion years ago, including microscopic spherules dubbed "blueberries", which are rich in hematite (see Fig. 8.24). (Courtesy of NASA/JPL/ASU.)

during subsequent drier periods of the planet's history, but impact craters have exposed the clay minerals in thousands of locations across Mars. In at least one instance, rivers have apparently eroded the surface and formed deltas that indicate the sustained deposit of large amounts of clay in the vicinity of, and within, a pre-existing crater (Fig. 8.22). These delta clay deposits provide evidence for the long-term flow of liquid water in the Nili Fossae region of Mars. Minerals deposited on a small volcanic dome rising from a shallow bowl named Nili Patera in the Syrtis Major volcanic region of equatorial Mars suggest the presence of heated water, or hot springs, about 3 billion years ago.

Instruments aboard the *Mars Reconnaissance Orbiter* have detected other hydrated, or water-containing, mineral deposits spread over large regions of Mars between 2 and 3 billion years ago. They included hydrated silica, commonly known as opal, and hydrated sulfate, formed from the evaporation of salty and acidic water.

Carbonate minerals have also been detected. They can be created when water and atmospheric carbon dioxide react in the presence of dust or volcanic rock. The carbonates found so far are regional, not global deposits that might indicate an ancient Martian ocean.

2001 Mars Odyssey observations of ancient rocks containing olivine, which is easily destroyed by liquid water, nevertheless indicate that many areas of Mars have been dry for a very long time, and this has been confirmed by roving spacecraft on Mars.

Rovers obtain evidence for a watery past on Mars

The *Mars Pathfinder* lander was parachuted to Mars on 4 July 1997, in celebration of the United States Independence Day. It was directed to Chryse Planitia, the Plains of Gold, chosen because it lies at the mouth of the large outflow channel Ares Vallis and at low elevations where

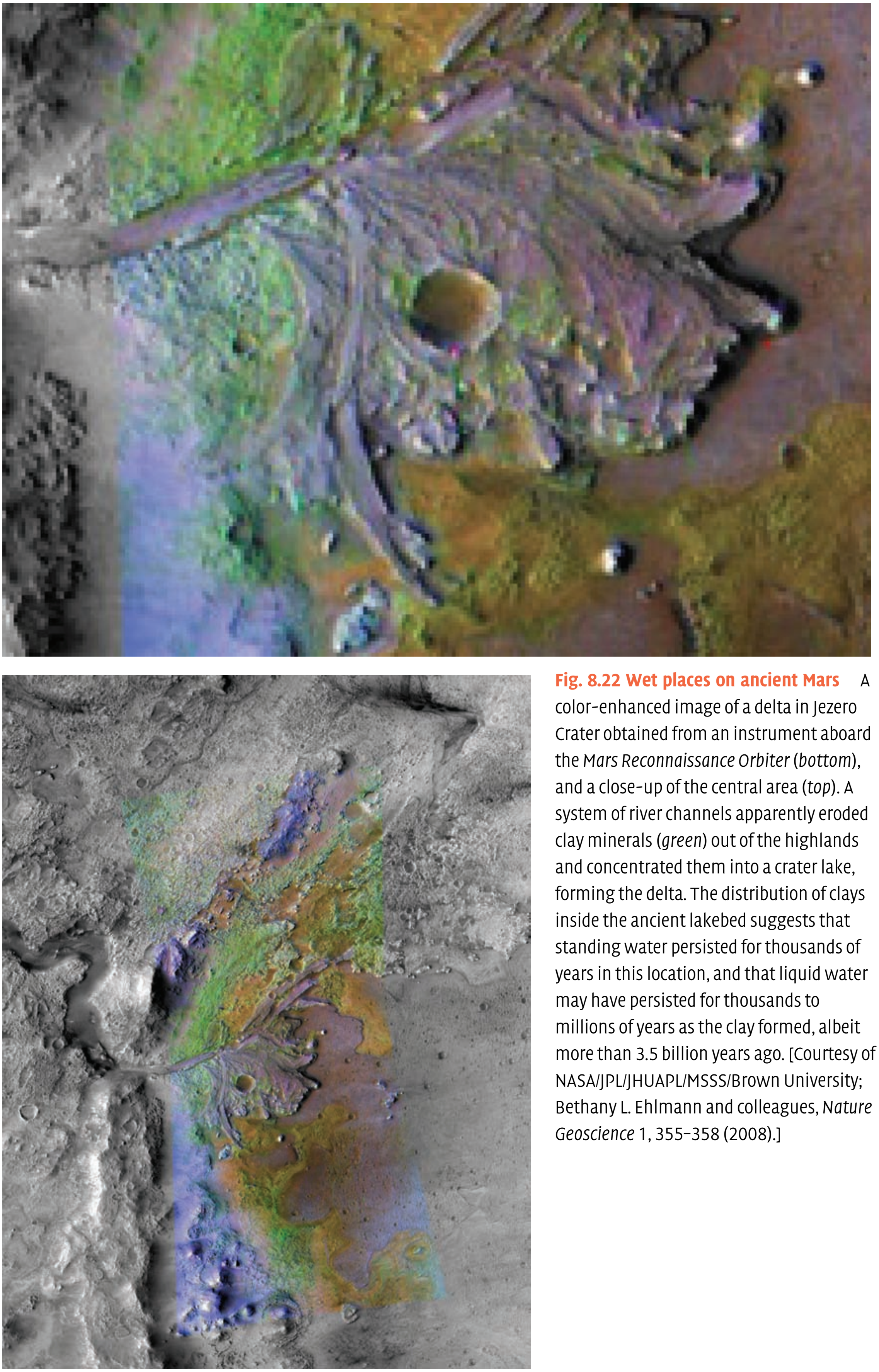


Fig. 8.22 Wet places on ancient Mars A color-enhanced image of a delta in Jezero Crater obtained from an instrument aboard the *Mars Reconnaissance Orbiter* (*bottom*), and a close-up of the central area (*top*). A system of river channels apparently eroded clay minerals (*green*) out of the highlands and concentrated them into a crater lake, forming the delta. The distribution of clays inside the ancient lakebed suggests that standing water persisted for thousands of years in this location, and that liquid water may have persisted for thousands to millions of years as the clay formed, albeit more than 3.5 billion years ago. [Courtesy of NASA/JPL/JHUAPL/MSSS/Brown University; Bethany L. Ehlmann and colleagues, *Nature Geoscience* 1, 355–358 (2008).]

water might have accumulated. The spacecraft's *Sojourner Rover* therefore moved about the surrounding terrain in search for signs of water.

The size distribution and composition of the many rocks and boulders surrounding *Pathfinder* are consistent with their being deposited there by flowing water. In addition, the presence of numerous rounded pebbles implied the erosive action of running water in the past. The immediate vicinity of the landing site nevertheless appears to be dry and unaltered since catastrophic floods sent rocks tumbling across the plain more than 2 billion years ago. It has apparently been untouched by water ever since the ancient deluge.

The two *Mars Exploration Rovers*, *Spirit* and *Opportunity*, were deployed in January 2004, roaming the surface of Mars for more than five years. The *Spirit* rover was placed within Gusev Crater, an ancient impact crater whose smooth flat floor was interpreted as sediments deposited in a crater lake (Fig. 8.23). The water network Ma'adim Vallis cuts through the southern rim of Gusev Crater, apparently draining the ancient cratered highlands to the south. *Opportunity* was sent to Meridiani Planum, following the discovery of hematite there, an iron-oxide mineral that typically forms in the presence of liquid water. This mineral had previously been found in the region using a spectrometer aboard the *Mars Global Surveyor* orbiter.

As expected, *Opportunity* found that Meridiani has a water-rich history, gathering compelling chemical and mineral evidence that this region of Mars stayed wet for an extended period of time long ago. There were the "blueberries", formed by slow evaporation in mineral-rich liquid water (Fig. 8.24), and fossilized ripples in nearby sedimentary rock, attributed to the sloshing of shallow-water waves. The spherical blueberries are rich in the iron-bearing mineral hematite, explaining the mineral signatures seen by the *Mars Global Surveyor* orbiting spacecraft. When *Spirit* became stuck in 2009, its churning wheels broke through the crust, like the spinning wheels of a car stuck on an icy road, revealing former subsurface water flow in soil layers now covered by wind-blown sand and dust. It turned out, however, that the minerals at the landing site could only have precipitated from highly acidic water, more like battery acid than drinking water.

When the floor of Gusev Crater turned out to be just a rock-strewn plain with no water in sight, the disappointed scientists directed *Spirit* to the nearby Columbia Hills where they found what they were looking for. The hills have been extensively altered by flowing water.

So *Opportunity* and *Spirit* have provided evidence that opposite sides of Mars were once soaking wet. Nevertheless the Meridiani Planum evaporates and the Columbia Hills rocks only indicate a wet environment before 3.7 billion years ago when the conditions were also acidic, oxidizing and very salty. Since then the regions have dried out and frozen into ice, modified by winds, dust and impacts rather than flowing water.

Fig. 8.23 Mars' Ma'adim Vallis and Gusev Crater A mosaic of *Viking* images reveals the branching, winding valley network named Ma'adim Vallis, which is about 800 kilometers long. It drains from the heavily cratered terrain in the south (*bottom*) and breaches the southern rim of the Gusev Crater (*top*), which formed earlier and is about 150 kilometers in diameter. The smooth flat crater floor suggested it was filled with water and sediments, but the *Mars Exploration Rover*, *Spirit*, failed to find any water-related sediments in the hypothetical crater lake, perhaps because they have been covered by volcanic flows. (Courtesy of NASA/JPL/USGS.)

Where did all the water on Mars come from and where did it all go?

There are several possible explanations for the origin of the water networks between 3.5 and 3.8 billion years ago. The water could have fallen as rain, during a sustained period of warm, wet climate, or it might have flowed just below the Martian surface, warmed by internal heat. According to another theory, the bombardment of comets and asteroids

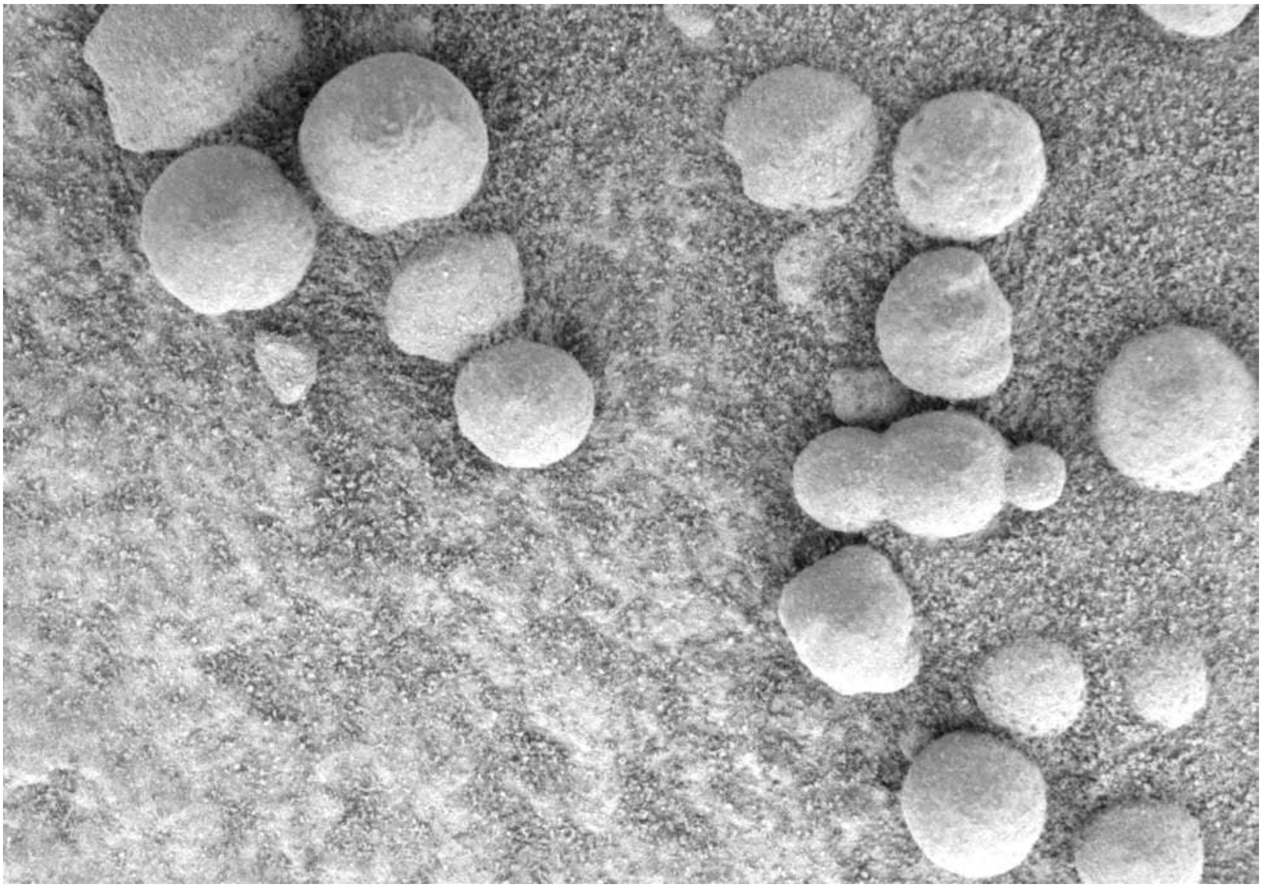


Fig. 8.24 Martian blueberries This microscopic image taken from *Opportunity* shows sphere-like grains dubbed "blueberries" that fill a depression near the landing site of the *Mars Exploration Rover* in Meridiani Planitia. The spherical structures are concentrations of the mineral hematite formed in liquid water about 3.7 billion years ago. (Courtesy of NASA/JPL/Cornell U./USGS.)

on early Mars may have caused torrential rains and the formation of the ancient networks.

Where did the floods that formed the outflow channels come from? All of the channels emerge from discrete sources in areas that have undergone collapse, suggesting that the rapid melting of subsurface ice filled the outflow channels with raging floods. Three of the largest outflow channels, the Ares, Simud and Tiu Valles, originate in the chaotic terrain, regions of fractured, jumbled rocks that apparently collapsed when groundwater suddenly poured out.

What triggered the sudden release of such huge volumes of water in the outflow channels? Liquid water might have been trapped beneath a thick frozen expanse of permafrost. When volcanic eruptions or the formation of an impact crater breached the overlying frozen seal, the underground water would be suddenly released under great pressure. The rapid surge of water would create dramatic and sudden floods, each lasting only a few days, weeks or months. The overlying surface layer would then collapse, creating chaotic terrain. In an alternative explanation for the outflow channels, Mars may have once been warmer and wetter, wrapped in a thicker atmosphere than it is now, with rain falling in torrents from the sky and coursing across the ground.

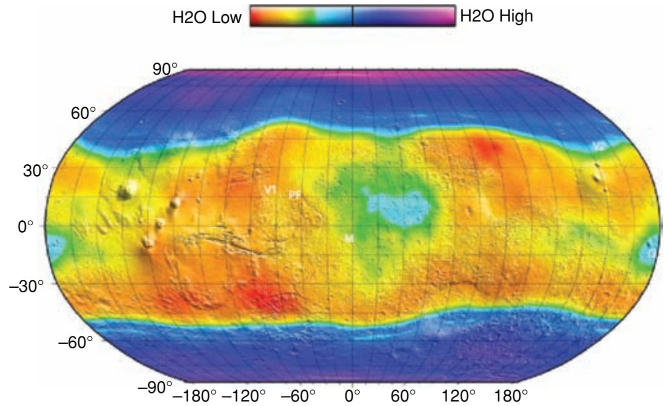
Everyone agrees that water once flowed on the surface of Mars in large quantities, and that much of that water has now left the surface and atmosphere. There is no liquid water residing on the surface of Mars today, and the amount of water vapor in its atmosphere is negligible. The Martian atmosphere might have once been warm and dense enough to allow water to remain liquid on or near the surface of Mars, possibly with rainfall. But then something changed. The atmosphere nearly disappeared and most of the water turned into ice.

8.8 Mars is an ice planet

Crater ejecta suggest subsurface ice on Mars

Although the round shapes of impact craters on Mars resemble those on the Moon, the ejecta blankets

Fig. 8.25 Yuty Crater on Mars The rounded, layered ejected material surrounding this crater may have been created when an impacting object melted the permafrost, or frozen ground, on Mars. Multiple layers of successive flows resemble the overlapping petals of a flower. The thin flow partly buries one crater, and is halted and deflected by the rim of another one. The muddy sludge seems to have sloshed across the surface, and was then refrozen. Such ejected features have not been found around craters on the Earth's Moon or the other planets. This crater, named Yuty after a town is Paraguay, is 19.9 kilometers in diameters and located in Chryse Planitia at 22.4 degrees north and 34.2 degrees west. (A *Viking* image courtesy of NASA.)



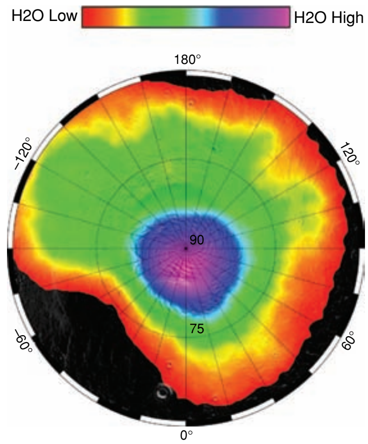


Fig. 8.26 Hydrogen on Mars A gamma-ray spectrometer aboard the *2001 Mars Odyssey* orbiter has provided a global map of the element hydrogen (*top*). Regions of high hydrogen content at the north and south polar regions (*dark blue and violet*) are attributed to very high concentrations of buried water ice, at well over 50 percent by volume. The equatorial regions of Mars (*red and yellow*) are significantly drier than the polar regions, although they do exhibit low concentrations of hydrogen. The water-ice region near the north pole (*bottom*) is due to a permanent polar cap of water ice on the surface; elsewhere in the north polar region abundant buried subsurface ice is also found. (Courtesy of NASA/LPL/U. Arizona.)

surrounding many Martian craters look quite different from their lunar counterparts. Craters on the Moon are surrounded by secondary craters, which were formed by material thrown out by the initial impact, just as you would expect from an explosion on the dry surface. In contrast, the material around fresh-appearing craters on Mars, especially those between 5 and 100 kilometers across, flows out in rounded fronts, each outlined by a ridge, somewhat like the splashed pattern formed when a pebble is dropped in mud (Fig. 8.25).

The flowing pattern of the Mars ejecta can be explained by supposing that the Martian ground contained water ice when they were formed. The ice would have extended a few meters below the surface, somewhat like the layers of permafrost underlying the Arctic landscapes on Earth. The heat of the explosive impact would have melted the ice, and the resulting steam and liquid water would then act as a lubricant for the flowing debris. The muddy material would have sloshed outward like a wave until it dried and stiffened, or became cool and refroze.

An instrument aboard the *Mars Reconnaissance Orbiter* has identified several locations at northern mid-latitudes where bright material surrounds fresh, relatively recent craters. This material is attributed to water ice excavated from beneath the surface at depths of up to 10 meters.

Vast amounts of subsurface water ice on Mars

More direct evidence for subsurface ice has been provided by the gamma-ray and neutron spectrometers aboard the orbiting *2001 Mars Odyssey* spacecraft. They detected strong signals from hydrogen, which is attributed to extensive water-ice deposits up to a meter below the surface and all over the planet (Fig. 8.26). The amount of hydrogen in the upper meter rises to above 50 percent by mass in vast regions surrounding both poles; it makes up as much as 10 percent by mass in the top meter of material in some regions close to the equator. The concentration of hydrogen is so large that it can only be due to water ice.

The ground-penetrating radar aboard the *Mars Reconnaissance Orbiter* has additionally found evidence for thick masses of buried ice extending for hundreds of kilometers across the mid-latitude region of northern Mars, outside the polar regions and close to the equator. The hidden glaciers and ice-filled valleys are located beneath coverings of rubble and debris. They may be the protected remnants of a former ice sheet that retreated when the climate changed. Bright, exposed ice excavated from fresh craters provides additional evidence for water ice hidden just below the surface of mid-latitude Mars.

The various measurements of subsurface water ice cannot address how much water could be present below the one-meter depth to which the techniques penetrate. But the huge amounts of water ice detected in the outer parts of the planet are not inconsistent with a thick, cold reservoir of water that extends kilometers deep. Closer to the surface, a robotic spacecraft has dug a trench into the soil at its landing place in the north polar plains of Mars, detecting water ice.

Phoenix lands on Mars

The *Phoenix* lander parachuted down to a northern latitude of about 68 degrees on 25 May 2008 and verified the presence of subsurface water ice there. Its robotic arm dug into the Martian soil (Fig. 8.27), and returned it to a bake-and-sniff oven that identified water vapor produced by heating the soil. A shallow layer of water ice was uncovered at depths of 5 to 18 centimeters.

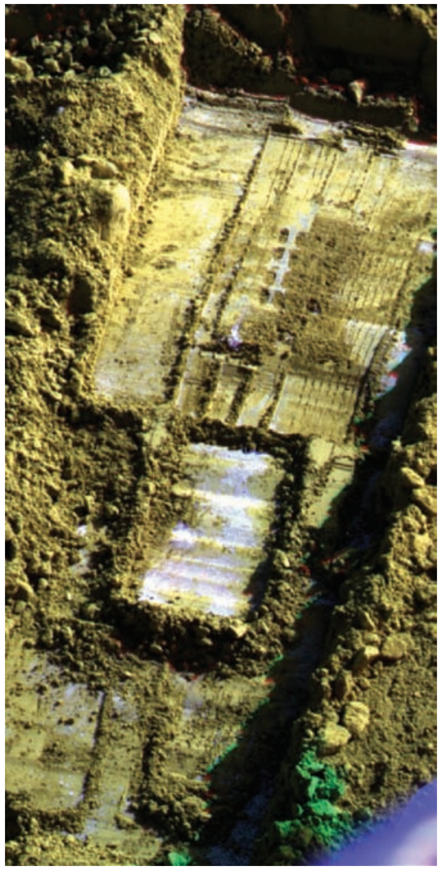


Fig. 8.27 *Phoenix* trench in the northern plains of Mars Bright, subsurface water ice has been exposed in this trench dug by the *Phoenix* lander in the northern Martian plains. The false-color image enhances the visibility of morning frost, especially around the edges of the trench, which are 4 to 5 centimeters deep and about 23 centimeters long. The details and patterns surrounding the *Phoenix* lander showed an ice-dominated terrain as far as it could view. (Courtesy of NASA/JPL/U. Arizona.)

The *Phoenix* soil experiments also provided evidence that liquid water has interacted with the Martian surface throughout the planet's history into modern times, and that volcanic activity has persisted into geologically recent times, several million years ago. It is nevertheless freezing cold everywhere on Mars in modern times. An ingenious light detection and ranging instrument aboard *Phoenix* even detected snow falling from water-ice clouds just above the landing site. In other places, it looks as if liquid water might have been released from under the frozen surface of Mars, but most likely not recently.

Gullies on Mars

Close-up images taken from the *Mars Global Surveyor* have revealed gullies that suggest water might have flowed in brief spurts on Mars in the relatively recent past. Tens of thousands of the gullies are found on the steep inside walls of craters or the crests of sand dunes, with forms that resemble gully washes on Earth. The Martian flow features emerge high up on the wall, run downhill in deep, winding channels, and fan out with an abrupt ending in a fan-shaped apron of dirt and rock (Fig. 8.28).

The gully flows cut through terrain that is relatively young, including sand dunes and crater-free landscapes. Unlike the immense outflow channels gouged out of the terrain billions of years ago, the lack of small craters superimposed on the small, unassuming gullies and their debris indicates that they are no more than a few million years old, and possibly a lot younger; older regions are pocked with craters. Some of the gullies appear to be dust-free, which may imply that they happened within the last year or two. Otherwise, the planet's perennial dust storms would have partly covered or even buried them in dust.

A few fresh-looking gullies were found in *Mars Global Surveyor* images taken in 2004 and 2005, which were not present in images taken a few years previously, leading to the controversial speculation that the gullies might be forming today, from liquid water that is now present under the surface of Mars.

Fig. 8.28 Gullies with possible water-carved channels on Mars This false-color image, taken from a high-resolution camera aboard the *Mars Reconnaissance Orbiter*, shows gully channels in a crater in the southern highlands of Mars. The gullies emanating from the rocky cliffs near the crater rim (*upper left*) show meandering and braided patterns characteristic of water-carved channels. These gullies offer strong evidence of liquid water flowing on Mars, perhaps within the last few million years, whereas different, fresh gully deposits in other regions may have resulted from landslides of loose, dry material. North is approximately up and the image width is about 1 kilometer. (Courtesy of NASA/JPL/U. Arizona.)

When the Martian surface was scrutinized in greater detail, the planet nevertheless looked much drier than had been previously supposed. Features examined with high-resolution instruments aboard the *Mars Reconnaissance Orbiter* indicate that the bright-looking deposits of new gullies are most likely landslides of sand or loose dust, and that fresh flows on some sand dunes may be caused by carbon-dioxide frost rather than water.

Images of supposed ancient ocean floors and riverbeds also show no obvious signs that liquid water was ever present, suggesting that they are now buried beneath dust or lava. This might explain why spacecraft that landed on putative Martian lakebeds and outflow channels found no evidence for recent water flow.

There are some gullies, however, that still offer strong evidence of liquid water flowing on Mars within the last few million years. These gullies are on slopes too shallow for dry flows and the high-resolution *Reconnaissance* images show clear indications of liquid flows, such as meandering and braided channels and terraces within the gullies.

A peculiar thing about both the dry and wet gullies is their location. Nearly all of them occur in places that are well below freezing all year round. Like moss on trees, the gullies form on the coldest slopes facing away from the Sun, where they are usually in shadow and rarely warmed by sunlight. Snow and ice would most likely now accumulate in these places, but when the polar regions tilted toward the Sun they would obtain enough sunlight to warm the topmost layer above freezing. This would explain why the Martian gullies are found at high latitudes and on poleward-facing slopes.

In this theory, gully formation isn't happening right now but when the swaying tilt of Mars' rotation axis is greatest. The poles then get more Sun exposure, the snow melts and gullies are formed. This tilt is measured by the planet's obliquity – the angle between the planet's axis of rotation and a line perpendicular to the plane of its orbit. Calculations indicate that the obliquity oscillates back and forth between 15 and 35 degrees every 100 000 years, with wider excursions occurring every 1 to 10 million years. The planet's rotation axis is currently tilted at 25.2 degrees, and when it becomes more than about 30 degrees the top layer of the Martian permafrost will thaw. Gullies might then be created by water mixed with ice.

The ongoing hunt for water on Mars is part of NASA's recent "follow-the-water" approach to the search for life on Mars. Liquid water is a key ingredient of life as we know it, so where there is liquid water there might be life. Once you find it on Mars you at least know that this region of the planet is habitable, either now or in the remote past depending on when the water flowed. Even if the water flowed millions or billions of years ago, hardy microbes might have survived until now, or perhaps there has never been a living thing on Mars. It's all speculation, until we land on the planet and find out. That is exactly what NASA is preparing to do, hoping to send robotic spacecraft to the sites of past water flow in the search for past or present life.

8.9 The search for life on Mars

Did any living things originate and survive on Mars?

Could life originate on Mars? No one knows for sure. There is abundant evidence for flowing water in the distant past, a few billion years ago, but not all water is fit to drink. The acidic, salty water that might have once flowed on Mars could have thwarted the chances of even microbes from developing or surviving.

Also, Mars subsequently lost its magnetic field and most of its atmosphere, making the Martian surface an extremely hostile place by Earth standards. Today Mars does not have enough oxygen, liquid water or heat for most, or possibly any, forms of terrestrial life. Its atmosphere is nearly all carbon dioxide, with very little oxygen or water vapor, and it is 150 times thinner than our air.

The lack of definite proof for liquid water on Mars today dims the chance for life there now. The red planet also now has no protective global dipolar magnetic field to keep potentially lethal solar energetic particles or cosmic rays from reaching the ground.

Moreover, without a protective ozone layer, the frozen surface is bathed in intense ultraviolet sunlight during the daytime, at levels that would quickly kill most organisms found on Earth today. The radiation damage to cells by the Sun's ultraviolet rays is indeed a serious risk to astronauts traveling to Mars. Any human visitor to the planet would have to wear a spacesuit to provide protection from the Sun's harmful radiation, as well as supplying oxygen to breathe.

In addition, the planet's soil now contains powerful oxidants that can break apart organic molecules, which provide living things with the capacity to evolve, adapt and replicate. Many scientists therefore remain skeptical about the chances of now locating living organisms in the dry, cold, hostile Martian world.

Optimists, on the other hand, have thought of plausible ways that primitive Martian life might survive when harsh conditions arrived. They could have taken refuge under the frozen Martian surface, for example, where it could be warmer and liquid water might be found. They might even be extracting water from the ice by chemical processes, developing internal antifreeze and hiding under or within rocks to avoid lethal ultraviolet sunlight and energetic solar and cosmic particles. After all, the optimists argue, single-celled life dominated the surface of the Earth for nearly 3 billion years, well before the rise of oxygen in its atmosphere, and microbial organisms now thrive on Earth under extreme conditions that were once thought to be lethal, including freezing temperatures underneath and within Antarctica sea-ice. All of these extreme life forms require water to survive, so the search for life on Mars ought to begin at places there might have been water. Even so, to send a spacecraft to Mars in search of life was an exciting long-shot gamble.

Viking 1 and 2 do not find evidence for life on Mars

One of humanity's most daring and imaginative experiments involved landing spacecraft on the surface of Mars, and searching for evidence of life there. The 4-billion-dollar gamble, in today's dollars, began on 20 July 1976, when the *Viking 1* lander came to rest on the western slopes of Chryse Planitia, the Plains of Gold, region of Mars. It appeared to have once been inundated by a great flood and was thus a promising place for life to have arisen. Six weeks later, the *Viking 2* lander settled down in the Utopia Planitia region on the opposite side of the planet, near the maximum extent of the north polar cap, again a favorable site for water and possible life.

Fig. 8.29 Apparently lifeless surface of Mars The Martian surface in western Chryse Planitia, as viewed from the *Viking 1* lander on 3 August 1976. Wind-blown dust clings against the eroded rocks, creates dust drifts and fills the sky. These drifts were little changed during the six years they were observed from *Viking 1*, and revealed no evidence for the movements of any living things. Mars is instead a cold and desolate world in which the silence is broken by the roar of winds, the hiss of dust, the rumble of mammoth landslides, and perhaps by outbursts of active volcanoes. (Courtesy of NASA/JPL.)

How did the *Viking* landers test for life on Mars? The first, most obvious test consisted of looking to see if any living thing was moving about on the Martian surface. Pictures were taken of all the visible landscape, from the stubby lander-legs to the horizon, for two complete Martian years, but the view was always one of a desolate, rock-strewn, wind-swept terrain (Fig. 8.29). A careful inspection of all these pictures failed to reveal any motions or shapes that could suggest life, not a single wiggle or a twitch, or an insect or worm. So there are probably no forms of life on Mars larger than a few millimeters in size.

Of course, no one really expected that the *Viking* eyes would see living things frolicking about on the Martian surface, and each of the landers carried a biology laboratory designed to measure samples of the Martian soil for organic molecules and for signs of growth that might signal the presence of living microorganisms.

The presence of microbes might be inferred if the Martian soil contained organic molecules, which are formed by combining carbon atoms with other atoms such as hydrogen. All living things on Earth contain organic molecules, but non-biological organic molecules have already been found in comets, meteorites, Titan's atmosphere, and in abundant quantities in the space between the stars. It is not at all certain that living things could have originated from any of these organic molecules; so finding them on Mars would not be a conclusive indicator of life.

A test for biological organic molecules in the Martian soil might be called a dead-body test, for soil would be expected to contain a higher proportion of organic molecules derived from dead bodies than from living ones. But the *Viking* experiments did not detect even a single carbon compound, even though the instruments could have spotted organic molecules at a concentration of one in a billion. The tests were so sensitive that they would have easily detected organic molecules from the most barren and desolate environments on Earth.

The other experiments on board the *Viking* landers searched for the vital signs of living microbes. They did this by exposing the soil to various nutrients, and sniffing the atmosphere to see if any hypothetical microbes ate the food and released gas. Something did emit carbon dioxide and oxygen gas, but it wasn't alive.

When samples of the Martian soil were exposed to liquid food laced with radioactive carbon, large amounts of radioactive carbon dioxide poured out from the soil. This certainly suggested that animal-like microbes were digesting the food and exhaling carbon dioxide gas. But when additional nutrients were added to the soil, there was no additional increase in radioactive gas. Living creatures would have continued to ingest the food.

When the Martian soil was exposed to water, a burst of oxygen flowed from it. At first, the surprised scientists thought that plant-like microbes were emitting the oxygen, but they soon realized that the release was too fast and brief. Microbes would grow and produce more oxygen as time went on, releasing oxygen at a steady rate. Moreover, the oxygen was detected when the experiment was performed in the dark, and this behavior would not be expected from plant-like photosynthesis that depends on sunlight.

After further experiments, scientists concluded that the biological tests failed to detect any unambiguous evidence for life on Mars. Instead of being produced by organisms of any kind, all of the results were attributed to non-biological, chemical interactions. Highly oxidized minerals in the Martian soil were reacting with the nutrients, breaking them up and liberating some oxygen gas and even more carbon dioxide. Thus, at present we have no irrefutable evidence for even microbial life now on the surface of Mars.

The Sun's ultraviolet radiation has apparently turned the atmosphere and soil into an antiseptic, oxidized form. The lethal soil would destroy cells, if there were any in the first place, living, dormant or dead, wiping them out with chemical reactions. The highly oxidizing conditions have even colored Mars red, turning iron in its surface material into rusted iron oxide.

The pioneering *Vikings* paved the way for the next spacecraft to land on the reddish-brown surface. They included the *Sojourner Rover*, deployed from the *Mars Pathfinder*, and the two *Mars Exploration Rovers*, *Opportunity* and *Spirit*. They moved across the Martian terrain, unlike the two *Viking* landers that were confined to one location, but their cameras have also never detected any signs of moving creatures. But these rovers did not carry out any chemical tests for living microbes. As we have seen, they instead searched for signs of liquid water, which might suggest that Mars could have supported life, either in the remote past or relatively recently. In the meantime, a dramatic, widely publicized frenzy resulted from speculative accounts of ancient microscopic life in a meteorite from Mars.

Possible life in a rock from Mars

Rocks that arrive from space and survive their fiery descent to the ground are given the name *meteorites*. Thousands of them have been recovered from the ice sheets of Antarctica, and most are chipped fragments of asteroids, a ring of rubble located between the orbits of Mars and Jupiter.

Only about a dozen meteorites have been identified as coming from Mars. They have been collectively named the SNC meteorites after the initials of the locations where they were first observed to fall from the sky – near Shergotty, India, in 1865, Nakhla, Egypt, in 1911, and Chassigny, France, in 1815. Pockets of gas trapped in the SNC meteorites have a unique composition that exactly matches that of the same gases in the Martian atmosphere, indicating that these rocks came from Mars and nowhere else.

The origin of each meteorite from Mars can be traced back to the jolt of a much larger object that struck the planet long ago. Most of the debris of this violent collision would fall back to the Martian surface, but some of it would be blasted off at a high enough speed to escape the weak tug of Martian gravity. That material would move in its own orbits around the Sun. These orbits would be gradually skewed by the gravitational pull of the distant planets, and very occasionally redirected on a collision course with Earth. Over an interval of 10 to 100 million years, a small fraction of the ejected debris would eventually strike the Earth.

Interest in the possibility of Martian life was heightened when scientists found possible signs of ancient, primitive bacteria-like structures inside one of these meteorites. It was recovered from the Allan Hills region of Antarctica, the first to be processed from the 1984 expedition; hence the designation ALH (for Allan Hills) 84001.

Everyone agrees that ALH 84001 originated on Mars soon after the planet formed, and was found in Antarctica 16 million years after being blasted away from the red planet. The controversy and excitement centers on suggestions that ancient microbial life took refuge in cracks within the meteorite a very long time ago.

In 1996, there was a great deal of public excitement over reports that organic molecules and possible microfossils of bacteria-like organisms had been found in ALH 84001, suggesting the existence of ancient primitive life on Mars. Newspaper headlines reported that the signs of life had been found on Mars. But the putative life forms had to be fossils of long-dead corpses, living several billion years ago, and they could only be seen with the most powerful electron microscopes on Earth.

The evidence for ancient microscopic life in the meteorite from Mars has not withstood the test of time. After more than a decade of study, all the excitement has died away. Even further, controversial evidence for structures with similar size, shape and arrangement in another Martian meteorite has not convinced the skeptics. Most scientists currently prefer non-biological explanations for all the evidence. The so-called fossils are, after all, smaller than bacteria on Earth, and much smaller than a cell; they are most likely just some kind of mineral formation.

So the meteorites do not provide convincing evidence that life once existed on Mars, but they do suggest that conditions suitable for life might have been present on the red planet early in its history. Many scientists subscribe to the belief that life will be discovered in some location other than Earth, and it might be on Mars or some other place.

The continuing hunt for signs of life on Mars

The apparently negative results of the *Viking* biology experiments, combined with the current hostile environment on Mars, strongly suggest that the planet is now inhospitable to any sort of existing life, and scientists have generally stopped looking for it. They are instead trying to determine if life ever arose on the planet in the past, focusing on the possibility that it once existed when running water and a warm dense atmosphere might have resulted in more favorable conditions for life.

There is even some lingering hope for detecting the chemical signatures of Martian life. In the early 21st century, astronomers detected the spectral signatures of methane molecules in the atmosphere of Mars. Methane, the simplest of hydrocarbon molecules, with one carbon and four hydrogen atoms, is easily broken apart by ultraviolet light from the Sun, so it must have been put into the atmosphere within the past 300 years. There are two possible explanations for its existence. Geothermal chemical reactions, possibly involving underground water and heat, could be releasing the gas, or bacteria-like microorganisms currently living below the frozen Martian surface could produce the methane.

It seems that modern civilization has always anticipated finding life on Mars, perhaps because of its Earth-like seasons, clouds, past flowing water, ice caps and similar daily rhythm. And Mars remains the most likely, nearby place in the solar system to find extraterrestrial life. We have visited the Earth's Moon and concluded that there is no life there. The intense heat on the surface of Venus boiled away any water long ago; it would fry and vaporize all living things that we know of. Mercury has essentially no protective atmosphere, and its temperature extremes from day to night would alternately boil and freeze anything on its surface, except perhaps at the polar regions.

Thus, as dry and cold as it might be, Mars remains the most plausible nearby home for life in the solar system outside Earth itself, and future voyages to search for life on Mars are inevitable. NASA is, for example, now considering a *Mars Science Laboratory*, planned for launch as early as 2011, to search for life in the Martian soil. It will collect samples from the landing site and analyze them for organic compounds and environmental conditions that could have supported microbial life now or in the past. The onboard chemical analysis will include the possible identifications of proteins, amino acids and other complex molecules that are essential to life as we know it. The instruments will also be able to identify atmospheric gases associated with biological activity.

Robotic spacecraft, and eventually humans, will visit the most likely places to contain life on Mars, returning rocks and soil to be scrutinized in the terrestrial laboratory, and this brings up the possibility of exchanging microscopic life between the planets (Focus 8.2).

The discovery of life on Mars, even primitive life in the very distant past, would have profound implications. It would give us companionship in a vast and lonely Universe, and it would also be a little humbling. The discovery would raise the likelihood that life might be found elsewhere in the Universe as well, perhaps on one of the planets that surely exist in our Galaxy. Of course, the enduring idea of life on Mars could prove as illusory as the Martian canals, but even in that event many humans will still retain a passionate conviction that there must be life somewhere else in the Universe. So the quest will continue, and whatever happens the human spirit will remain as beautiful and glorious as ever.

Even if life within our solar system is only found on Earth, the discovery of planets around nearby stars other

Focus 8.2 Microbes from Earth and Mars

Terrestrial microbes will surely accompany astronauts to Mars, perhaps contaminating the planet and complicating the search for life. Conversely, rocks and soils brought back to Earth from Mars by a future space mission could be full of deadly microbes that might cause a global catastrophe on Earth. Visiting astronauts could conceivably get infected with an alien Martian plague, and be forced into quarantine if they make it back home. Precautions should also be taken against an unexpected crash landing of the returning spacecraft, which could release dangerous alien organisms. Since the surface of Mars is now sterilized by solar ultraviolet rays and the planet's oxidizing ground, these are unlikely possibilities, but precautions are being taken just in case the improbable occurs.

As a matter of fact, it is possible that Earth and Mars have regularly exchanged microbes over the years, without any modern spacecraft or astronauts being involved. Life forms may have arisen on Mars first and then migrated to Earth, hitching a ride on a meteorite; or it might have been the other way around, with life originating on Earth and traveling to Mars. Cosmic impacts with Mars have sent hundreds of tons of nomadic Martian rocks to Earth over recent centuries, and much more during the past eons. The rain of impacting debris was most intense in the early days of both planets, increasing the likelihood of biological exchange between them. And even now, two tons of Martian rocks are thought to rain down on Earth each year, and about the same amount of terrestrial rocks annually smashes into Mars.

So, life might not have emerged spontaneously on Earth. It could have come from Mars, and in that case we might all be Martians. Or life might have originated on Earth and was then delivered to Mars. Maybe life arose on the two planets independently and spontaneously, or perhaps impacting comets or asteroids pollinated both planets. And just maybe Earth is the only place to harbor life in the entire solar system. It's all a lot of fun to think about, but every one of these possibilities is mainly speculation with little hard scientific evidence.

than the Sun (Chapter 16) suggests that life might be common in the cosmos at large. These stars and their planets were formed in interstellar clouds that contain vast amounts of water, one of the key molecules of life, as well as all kinds of organic, carbon-bearing molecules, from formaldehyde to benzene. Many scientists think that somewhere out there, among the 100 billion stars in our Galaxy, there ought to be at least one habitable planet similar to Earth, swarming with organisms and perhaps with something more advanced than us.

8.10 The mysterious moons of Mars

Discovery and prediction of the moons of Mars

Mars has two little moons that are so dark and small that they remained unseen for centuries, even after the invention of the telescope. They were discovered in 1877 by Asaph Hall (1829–1907) using a 0.66-meter (26-inch) telescope at the United States Naval Observatory. He named the inner moon *Phobos*, the Greek word for "fear", and the outer one *Deimos*, Greek for "flight, panic or terror", after the attendants of the Greek god of war, Ares, in Homer's *Iliad*.

Johannes Kepler (1571–1630) suggested the possible existence of two Martian moons as early as 1610. Since Venus has no moons, the Earth has one, and Galileo had discovered four large moons orbiting Jupiter, it seemed logical to Kepler that Mars, with an intermediate orbit between Earth and Jupiter, would have two moons.

Then Jonathan Swift (1667–1745) endowed Mars with two fictional moons in his *Gulliver's Travels*, published in 1726. He placed them close to Mars, at 6 and 10 Martian radii, near to the 2.7 and 6.9 of Phobos and Deimos. Swift's prediction had to be a lucky guess, for there was no telescope at the time powerful enough to detect the two moons.

These are not the same kind of object as the Earth's large and spherical Moon. Both moons of Mars are tiny compared with their planet and they orbit very close to it. Phobos and Deimos are only a few tens of kilometers across and have insufficient gravity to mold them into a spherical shape (Table 8.7). They have an irregular shape and battered appearance, with craters large and small (Fig. 8.30), indicating that their surfaces are at least 3 billion years old. Both Martian moons move within the planet's equatorial plane, but they orbit so near to the surface of Mars that an observer at the poles of Mars could see neither moon. The two moons of Mars also have an unusually low mass density of less than 2000 kilograms per cubic meter.

A maverick moon of Mars

Phobos is the real maverick. It moves around Mars at 2.766 Martian radii, so close that it rises and sets three times in

Table 8.7 The satellites of Mars[a]

	Phobos	Deimos
Mass (kilograms)	1.065×10^{16}	1.47×10^{15}
Mean radius (kilometers)	11.1 ± 0.15	6.2 ± 0.18
Mean mass density (kilograms per cubic meter)	1872 ± 76	1471 ± 166
Mean distance from Mars[b]	$9378\,\text{km} = 2.766\ R_{\text{M}}$	$23\,479\,\text{km} = 6.926\ R_{\text{M}}$
Sidereal period (Martian days)[c]	0.3189	1.26244

[a] Adapted from JPL Planetary Satellite Physical Parameters.
[b] The mean radius of Mars is $R_{\text{M}} = 3389.5$ kilometers.
[c] One Martian day = 24 hours 37 minutes 22.663 seconds, so the orbital periods of Phobos and Deimos are 7 hours 39 minutes 13.84 seconds and 30 hours 17 minutes 54.87 seconds, respectively. Both satellites are locked into synchronous rotation with a rotation period equal to their orbital periods.

Fig. 8.30 Phobos and Deimos, moons of Mars A high-resolution camera on the *Mars Reconnaissance Orbiter* provided these close-up images of the Martian moons Phobos (*left*) and Deimos (*right*). The illuminated part of Phobos is about 21 kilometers across. The large crater Stickney, at the lower right, has a diameter of about 9 kilometers, nearly half the size of Phobos. This crater is named after Angeline Stickney (1830–1892), wife of Asaph Hall (1829–1907) who discovered the two small satellites of Mars. Deimos (*right*) is about 12 kilometers in diameter. (Courtesy of NASA/JPL-Caltech/U. Arizona.)

a single Martian day. That is, the orbital period of Phobos, of 7 hours and 39 minutes, is less than one-third of the planet's rotation period of 24 hours 37 minutes. From the surface of Mars, the small moon Phobos would be seen to move backwards across the sky, rising in the west and setting in the east.

Phobos is about as close to Mars as it can get. If it came much nearer to Mars, the planet's differential gravitational forces, between the sides nearest and furthest from Mars, would tear Phobos apart. In fact, the orbit of Phobos is steadily shrinking. If it continues to move toward Mars at the present rate, of about 1.8 meters closer to the planet every century, Phobos will either smash into the Martian surface or be ripped apart by the planet's gravity to make a ring around Mars in 50 million years. So astronomically speaking, we are now catching a fleeting glimpse of the last few moments of its life.

On the other hand, Deimos is near the outer limit for an object to be orbiting Mars. If it moved much further away, the Martian gravity would be too weak to hold on to the moon.

Phobos' suicidal motion has resulted in some fascinating, but untrue, speculations about the maverick moon. Two prominent scientists, the American astronomer Carl Sagan (1934–1996) and the Soviet astrophysicist Iosif Shklovsky (1916–1985), concluded in 1968, for example, that Phobos might be a hollow artificial satellite launched by a past Martian civilization; atmospheric friction, or air drag, would be causing the hypothetical satellite to move toward the planet.

We now know that tidal forces are pulling the small moon toward unavoidable destruction. Phobos produces two tidal bulges in the solid body of Mars, in much the same way that the Moon produces ocean tides on the Earth. As Phobos moves ahead of Mars, the closest tidal bulge pulls gravitationally on the moon, causing it to lose energy and move inexorably toward self-destruction. Because Phobos orbits so close and swiftly,

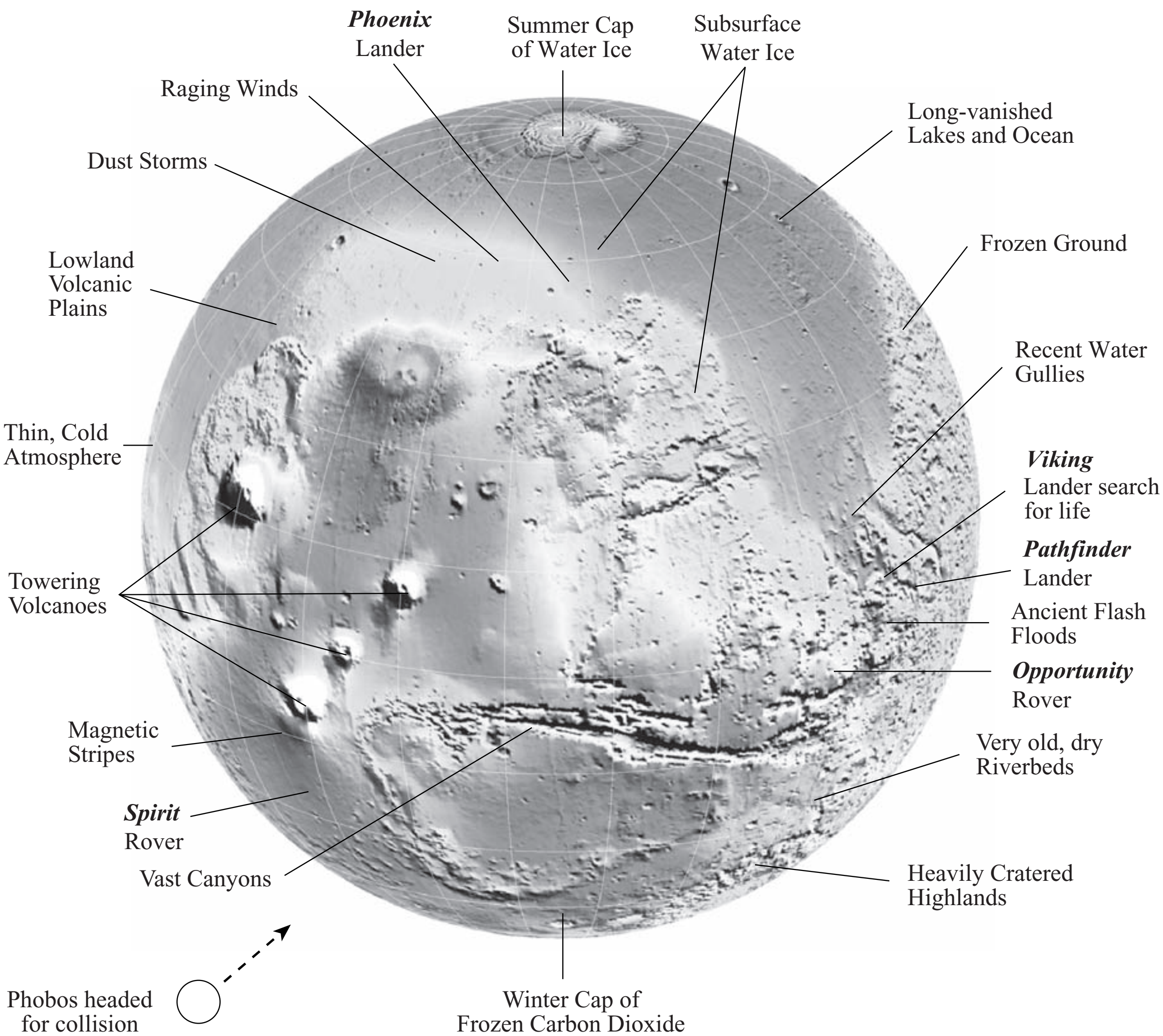


Fig. 8.31 Summary diagram

this tidal action pulls the moon inward, instead of pushing it outward as tidal interaction does for the Earth's Moon.

Origin of the Martian moons

The irregular shapes, small sizes, and low mass densities of Phobos and Deimos closely resemble those of the numerous asteroids that orbit the Sun between the orbits of Mars and Jupiter. The two Martian moons and the asteroids have a similar battered appearance, with a profusion of craters, large and small. Moreover, the composition of the Martian moons seems to be unlike that of the planet they orbit. Their surfaces are as dark as some asteroids, known as the carbonaceous C-type, and nowhere near as lightly colored as the surface of Mars. Scientists therefore speculate that Phobos and Deimos were adopted from the asteroid belt.

One idea is that two asteroids wandered close by Mars and were pulled in by its gravity, perhaps at different times. In another variant on this theme, one larger asteroid was captured by Mars, and subsequently broke apart during a collision with another larger body, becoming Phobos and Deimos. In both explanations, the captured asteroids would have to lose energy during their encounter with Mars. Otherwise they would hurtle back into space rather than going into orbit around the planet.

They might be the last surviving remnants of a ring of debris blasted into orbit by a huge meteorite that collided with Mars during its formative years, about 4 billion years ago. Such a giant impact is now the favorite explanation for the origin of the Earth's Moon. It would provide a good explanation for the nearly circular orbits of Phobos and Deimos, which lie in the equatorial plane of Mars; such orbits are difficult to explain by the captured-asteroid hypothesis.

Part 3 The giant planets, their satellites and their rings: worlds of liquid, ice and gas

9 Jupiter: a giant primitive planet

- All we can see on Jupiter is clouds, swept into parallel bands of bright zones and dark belts by the planet's rapid rotation and counter-flowing, east–west winds.
- Jupiter turns to liquid under high pressures within its interior, so the cloudy atmosphere has no distinct bottom and Jupiter's weather pattern is free to flow in response to the giant planet's rapid spin.
- Jupiter's Great Red Spot and white ovals are huge shallow anticyclonic storms, which can have diameters larger than the Earth's and last for centuries.
- Large whirling storms on Jupiter gain energy by merging with, and engulfing, smaller eddies. The little storms obtain their energy from hotter, lower depths.
- White clouds of ammonia ice form in the coldest, outermost layers of Jupiter's atmosphere. Water clouds are expected to form at greater depths, and ammonium hydrosulfide clouds should condense between the water and ammonia clouds.
- All of the clouds on Jupiter ought to be white; their colors are attributed to an active chemistry that produces complex compounds in small amounts.
- Bolts of lightning illuminate deep, wet storm clouds on Jupiter.
- When the *Galileo* spacecraft parachuted a probe into Jupiter, the entry site, a region of downdraft, was missing the expected three layers of clouds and it was far drier and windier than anticipated.
- The fierce winds that give rise to Jupiter's banded appearance run deep, indicating that Jupiter's ever-changing weather patterns are driven mainly from within, by internal energy rather than by external sunlight.
- When compared to the outer layers of the Sun, the outermost atmosphere of Jupiter is slightly depleted in helium, and enriched in carbon, nitrogen and sulfur by a factor of about three.
- Jupiter is a primitive incandescent globe that radiates 1.67 times as much energy as it receives from the Sun, probably as heat left over from when the giant planet formed.
- Jupiter originated together with the Sun, and both the giant planet and the star are mainly composed of the lightest elements, hydrogen and helium.

- If Jupiter was about 80 times more massive, it could have become a star.
- Jupiter has a non-spherical shape with a perceptible bulge around its equatorial middle.
- The visible cloud tops and outer atmosphere of Jupiter form a very thin veneer that covers a vast global sea of liquid hydrogen.
- Most of Jupiter's interior consists of fluid metallic hydrogen formed under the extreme pressures that exist inside the planet.
- Jupiter probably has a dense, molten core with a mass that is less than or equal to 12 times that of the Earth.
- By re-creating extreme conditions like those inside Jupiter, modern laboratory experiments have compressed liquid hydrogen so that it becomes highly conductive like a metal.
- Jupiter's powerful magnetic field is generated by rotationally driven electrical currents inside its vast internal shell of liquid metallic hydrogen.
- The volcanoes on Jupiter's innermost large moon Io have turned the satellite inside out. It is the most volcanically active body in the solar system.
- Io's volcanoes emit plumes of sulfur dioxide gas that freeze onto the surface as a white frost.
- Volcanic vents on Io are filled with melted silicate rocks that are hotter than any place on any planet's surface, even Venus.
- Changing tidal forces squeeze Io's rocky interior in and out, making it molten inside and producing volcanoes.
- A vast current of 5 million amperes flows between the satellite Io and the poles of Jupiter, generating 2.5 trillion watts of power and producing aurora lights on both the satellite and the giant planet.
- Jupiter's magnetic field sweeps past Io, picking up a ton of sulfur and oxygen ions every second and directing them into a doughnut-shaped ring known as the plasma torus.
- There are no mountains or valleys on the bright, smooth, ice-covered surface of Jupiter's moon Europa; it has few impact craters, indicating a relatively young age.
- Long, deep fractures run like veins through Europa's icy covering, apparently filled by the upwelling of dirty liquid water or soft ice. Warmer, slushy material just beneath the crust also lubricates large blocks of ice that float like rafts across Europa's surface.
- An electrically conducting, subsurface sea within Europa may be responding to Jupiter's magnetic field, generating a time-varying magnetism in the satellite.
- Scientists speculate that subsurface liquid water in Europa may harbor alien life that thrives in the dark.
- Jupiter's moon Ganymede is bigger than the planet Mercury. The satellite's icy surface has been fractured and pulled apart, producing a grooved terrain, and surface depressions have been filled by eruptions from volcanoes of ice.
- Ganymede has an intrinsic magnetic field, and it is the only satellite that now generates its own magnetism.
- Jupiter's moon Callisto has one of the oldest, most heavily cratered surfaces in the solar system. Yet the satellite is covered by fine, dark, mobile material and it has a lack of small craters when compared to the surfaces of the Moon and Mercury.

- Like Europa, the outermost large moon Callisto has a borrowed magnetic field, apparently generated by electrical currents in a subsurface ocean as Jupiter's powerful field sweeps by. But Callisto has a largely homogeneous interior without any apparent dense iron core, and the buried sea has to lie deep enough to not affect its unaltered, cratered surface.
- Jupiter's faint, insubstantial ring system is made of dust. The ring particles might last for only a few thousand years, and they must be replenished if the ring system is a permanent feature.
- When interplanetary meteoroids, attracted by Jupiter's powerful gravity, pound into the small inner moons of Jupiter, they chip off dust fragments that go into orbit around the planet, forming its ring system.

9.1 Fundamentals

Jupiter's exceptional brightness and stately motion among the stars earned it the reputation as king of the planets to ancient astronomers. The giant planet outshines everything in the night sky except the Moon and Venus, and it revolves around the Sun at a leisurely pace with an orbital period of 11.86 Earth years.

Jupiter's complete orbital journey across the background stars is close enough to 12 years that the Chinese adopted it for their 12-year astrological cycle, using the giant planet's motion to mark out the years. Named *Sui Xing*, for the "Year Star", Jupiter passes through patterns of stars representing an ordered sequence of a dozen animals. The arrival of the Chinese New Year at the end of January 2010 marked the beginning of the Year of the Tiger, and the succeeding years are designated as Rabbit, Dragon, Snake, Horse, Sheep, Monkey, Cock, Dog, Pig, Rat, and Ox. This system dates at least as far back as Marco Polo's (1254–1324) visit to the Mongol rulers of China. Jupiter also passes eastward through one of the twelve constellations of the Greek zodiac each year, but these stellar configurations are not related to the Chinese menagerie.

Jupiter's orbital radius is 5.2 times the radius of the Earth's orbit, so the planet's distance from Earth changes relatively little in the course of a year. As a consequence, its apparent size and brightness are fairly constant, unlike the behavior of Mars and Venus. When these nearby planets

Table 9.1 Physical properties of Jupiter[a]

Mass	1.8981×10^{27} kilograms = 317.894 M_E
Equatorial radius at one bar	71 492 kilometers = 11.19 R_E
Mean radius	69 911 kilometers
Bulk density	1326 kilograms per cubic meter
Rotation period	9.9249 hours = 9 hours 55 minutes 29.7 seconds
Orbital period	11.8626 Earth years
Mean distance from Sun	7.7833×10^{11} meters = 5.203 AU
Age	4.6×10^9 years
Atmosphere	86.4 percent molecular hydrogen, 13.6 percent helium atoms
Energy balance	1.67 ± 0.08
Effective temperature	124.4 kelvin
Temperature at one bar level	165 kelvin
Central temperature	17 000 kelvin
Magnetic dipole moment	20 000 D_E
Equatorial magnetic field strength	4.28×10^{-9} tesla or 14.03 B_E

[a] The symbols M_E, R_E, D_E and B_E denote respectively the mass, radius, magnetic dipole moment, and magnetic field strength of the Earth. One bar is equal to the atmospheric pressure at sea level on Earth. The energy balance is the ratio of total radiated energy to the total energy absorbed from sunlight, and the effective temperature is the temperature of a black body that would radiate the same amount of energy per unit area.

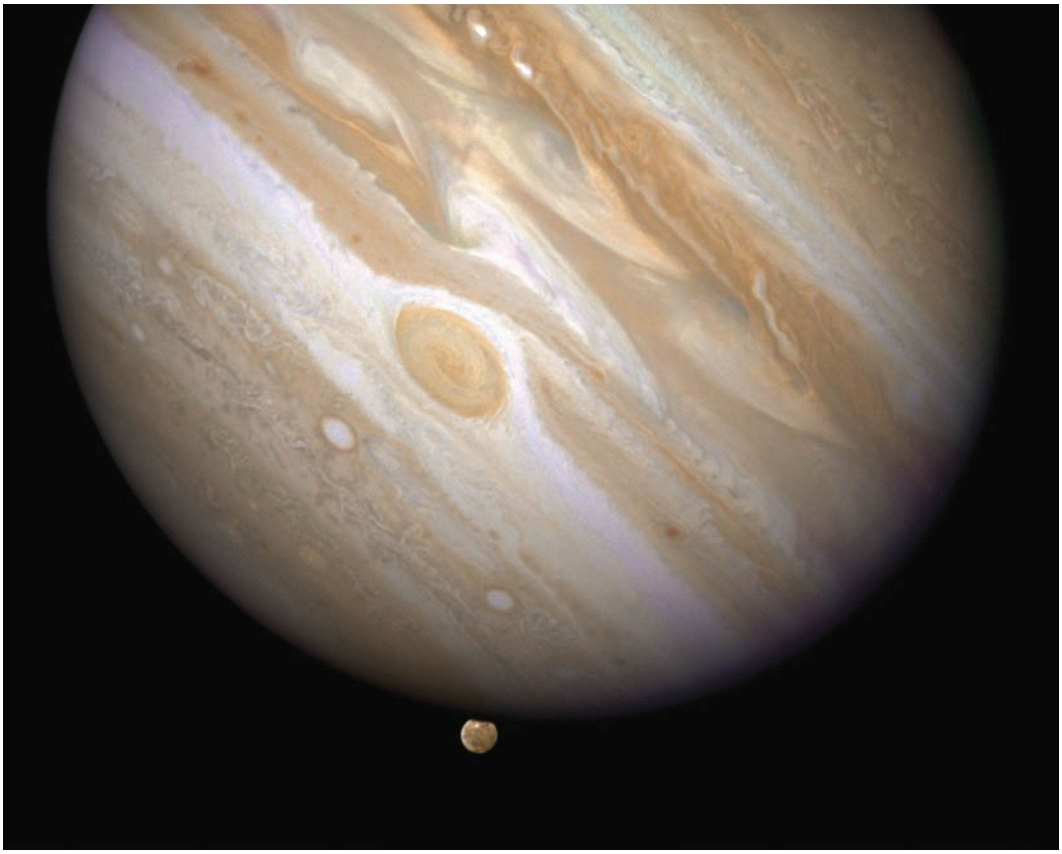


Fig. 9.1 Jupiter and its moon Ganymede The icy surface of Ganymede, Jupiter's largest moon, is seen just before it passes behind the planet (*bottom*). The giant Jupiter is portrayed in close to natural color, including its Great Red Spot (*center*), the biggest and oldest known storm in the solar system. The spot's east–west diameter is more than twice that of Earth, and one-sixth the diameter of Jupiter itself, and the red swirling vortex has been observed for at least 300 years. By watching Jupiter eclipsing its moon Ganymede, astronomers have searched for haze that causes a slight dimming at different colors. This image was taken on 9 April 2007 from the *Hubble Space Telescope*. (Courtesy NASA/ESA/E. Karkoschka, U. Arizona.)

are on the same side of the Sun as the Earth, they appear much bigger and brighter than when they move to the opposite side of the Sun.

Jupiter is a true monarch of the planets, the largest planet in the solar system with a radius of about 11 times that of the Earth. The giant is so large that it could contain more than 1300 Earth-sized planets inside its volume. Yet Jupiter is only 318 times as massive as our planet. So Jupiter must be mainly composed of something lighter than the rock and iron that constitute the Earth.

If we divide the mass by the volume, we find that Jupiter has a bulk density of 1326 kilograms per cubic meter, only about one-quarter the mean mass density of the Earth. In fact, the mass density of Jupiter is only slightly greater than that of water, at 1000 kilograms per cubic meter, and this implies that Jupiter, like the Sun, is composed primarily of hydrogen. No other element is light enough to account for the low density of the planet.

Despite its great size, Jupiter rotates so fast that day and night each last about 5 hours and its full day is less than one-half of an Earth day. The rapid rotation stretches most of Jupiter's storm fronts into strips around the entire planet. The precise rotation rate is found by tracking radio bursts that are linked to the planet's spinning magnetic field, which emerges from deep within the planet. A rotation period of 9 hours 55 minutes 29.7 seconds = 9.9249 hours is obtained from the repeated passage of the radio storm centers. This rapid rotation can easily be detected in an hour or so with a small telescope if the planet's cloud markings are carefully watched, but not measured with precision.

9.2 Stormy weather on Jupiter

Zones and belts in Jupiter's atmosphere

The only features we can see on Jupiter are multicolored clouds, and occasionally a large moon that passes by (Fig. 9.1). The clouds circulate around the planet in dark

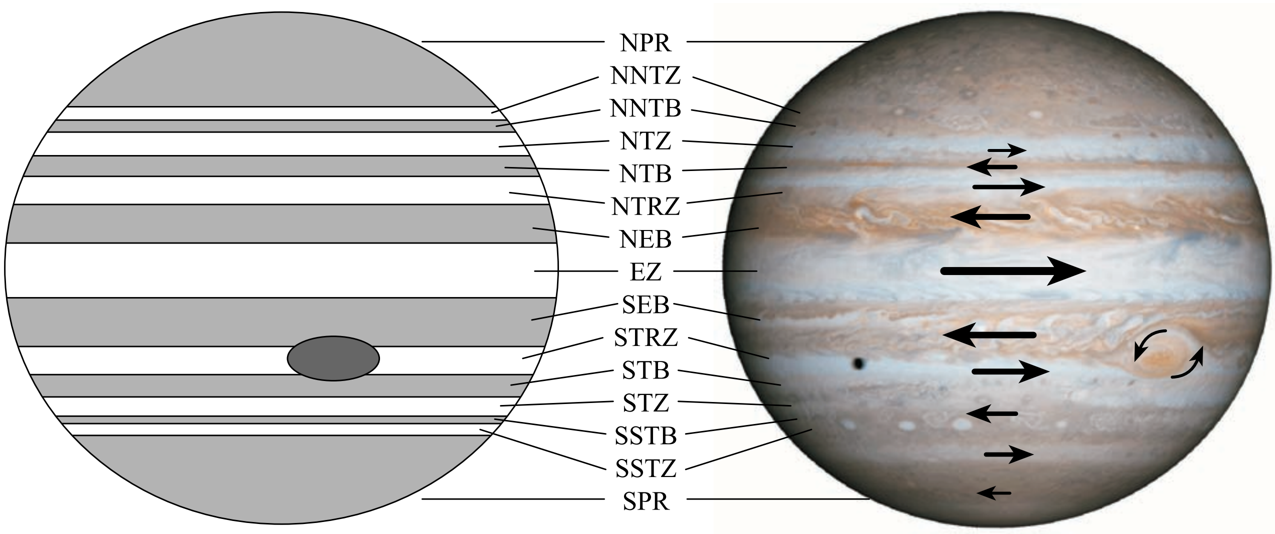


Fig. 9.2 Banded wind-blown clouds on Jupiter The traditional nomenclature of Jupiter's light and dark bands of clouds (*left*) is given in abbreviated form (*center*). The dark bands are called belts, denoted by "B", the light bands are known as zones, or "Z", and the rest of each name is based on climatic regions at the corresponding latitudes on Earth. North, letter "N" is at the top, and south, denoted by "S", is at the bottom. The equatorial, or "E", bands are in the middle, the tropical, "TR", bands on each side of the equator, and the temperate, "T", ones at mid-latitudes. Far northern latitudes are denoted by "NN", far southern latitudes by "SS", and the polar regions by "P". The image of Jupiter (*right*) was taken from the *Cassini* spacecraft on 7 December 2000, when Jupiter's moon Europa cast a shadow on the planet. The arrows point in the direction of wind flow, and their length corresponds to the wind velocity, which can reach 180 meters per second near the equator. (*Cassini* image courtesy of NASA/JPL.)

belts and light zones. They are the sites of counter-flowing winds, or jets, that blow parallel to the equator at different speeds and in different directions (Fig. 9.2).

The zones and belts are thought to be associated with vertical, in-and-out motions called convection. Upwelling warm gas results in the light-colored zones, which are regions of high pressure. The darker belts overlie regions of lower pressure where cooler gas sinks back down into Jupiter's atmosphere. The zones and belts are therefore analogous to the high- and low-pressure systems that produce localized circulating storms on Earth, except Jupiter's rapid rotation has wrapped them all the way around the planet. Although this general convective pattern of the zones and belts was apparently supported by observations during the *Voyager* flybys of Jupiter, instruments aboard the *Cassini* spacecraft have also discovered large plumes of fast-rising gas scattered throughout the belts and absent in the zones.

Where the Earth has just one westward air current at low latitudes – a trade wind, and one nearly eastward current at mid-latitudes – a jet stream, at the cloud-top level Jupiter has five or six of these alternating jet streams in each hemisphere.

The cloud-top winds on Jupiter have higher speeds than the winds on Earth, and storms last longer on the giant planet. Its raging winds are powerful jets, moving at speeds of up to 180 meters per second, more than four times faster than any jet stream on Earth. And unlike most winds on our planet, the east–west winds on Jupiter are remarkably steady. They vary in strength and direction as a function of latitude with a regular pattern that apparently has remained unchanged for at least a century.

On Earth, heating by sunlight results in a large temperature difference between the poles and equator, which drives our winds and circulates the air. But Jupiter's pole and equator share about the same temperature, at least near the cloud tops, so its winds are not just due to solar heating. An internal heat source most likely drives Jupiter's turbulent weather system from below.

Jupiter's storm clouds

Small eddies and larger vortices interrupt the smooth profile of the belts and zones on Jupiter with turbulent red, white and brown spots. These whirlpools mark patterns of weather, as the clouds billow, churn and seethe. Huge storms larger than the Earth in size swirl across the planet, while smaller eddies chase each other, whirling and rolling about. Unlike storms on Earth, underlying landmasses do not break up the storms on Jupiter. The planet's deep, windy atmosphere overlies a liquid hydrogen ocean.

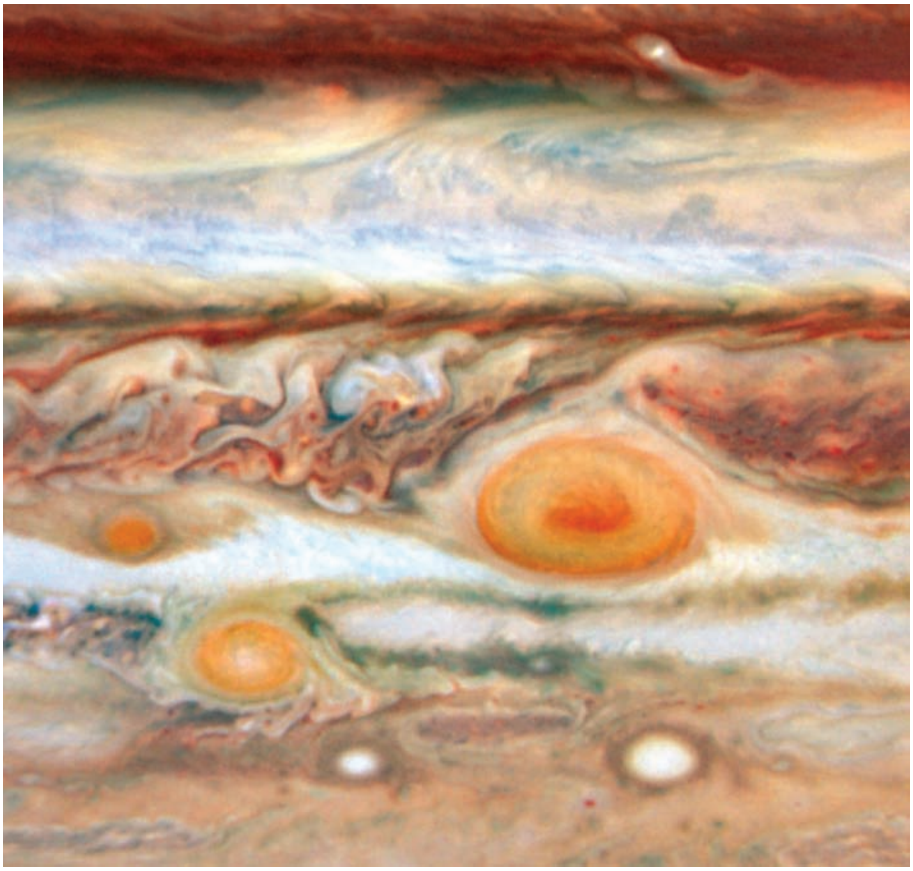


Fig. 9.3 New red spots on Jupiter Some of the turbulent eddies in Jupiter's atmosphere move from west to east (*left to right*) above Jupiter's Great Red Spot (*center*), a giant high-pressure anticyclone that swirls in the opposite counter-clockwise direction. Some small eddies are sucked into the great red vortex, helping to sustain it, while other eddies roll around the perimeter, probably reinforcing the storm's circulation. Although the Great Red Spot has been observed for at least three centuries, a new red spot appeared in the spring of 2006, and in 2008 a white oval-shaped storm turned into a third, smaller red spot. The small red oval (*left*) will either be absorbed or repelled by the Great Red Spot when it moves into it, but the other, lower latitude red spot, which lies between the other two, will most likely pass the Great Red Spot unhindered. This visible-light image was taken on 9–10 May 2008 from the *Hubble Space Telescope*. (Courtesy of NASA/ESA/Mike Wong and Imke de Pater, U. C. Berkeley.)

The biggest whirlpools are visible from Earth using small telescopes of just 0.08-meters (3-inches) aperture, and have been observed for centuries. The earliest sightings of a large spot in the atmosphere of Jupiter have been credited to Robert Hooke (1635–1702) in 1664 and to Giovanni Domenico Cassini (1625–1712) the following year. The most prominent one – the Great Red Spot – appears in records and drawings dating back to 1831, and might have coincided with the earlier sightings.

Jupiter's famous Great Red Spot is essentially a huge weather system, with an east–west dimension greater than Earth's diameter. Because of rapidly increasing pressure with depth it cannot extend deeply into the planet. It is simply an enormous, shallow eddy trapped between counter-flowing jets, so large that the strong prevailing winds are forced to flow around it. The winds, in turn, funnel smaller eddies toward the Red Spot, helping to energize it.

Winds are swirling inside the awesome vortex in the counter-clockwise direction, at speeds up to 110 meters per second. Since it is in the southern hemisphere of Jupiter, this rotational direction indicates that the Red Spot is a high-pressure vortex, known as an anticyclone. A low-pressure cyclone would spin in the opposite direction.

The Great Red Spot is not unique, but just the largest of hundreds of different storms on Jupiter, including egg-shaped white ovals and smaller red spots that come and go or merge together (Fig. 9.3). At least at cloud-top levels, most of these long-lived vortices are high-pressure anticyclones that rotate counter-clockwise in the planet's southern hemisphere and clockwise in the northern hemisphere, with counter-flowing winds on their sides. Instead of wandering unpredictably like terrestrial hurricanes, the titanic whirlpools on Jupiter drift at a steady rate in either the eastward or westward direction, apparently rolling between and with the winds, like a giant ball-bearing. In contrast, storms that rotate in the other cyclonic direction on Jupiter are short-lived, lasting several days or less before being torn apart by the action of shearing winds. Lightning observed by the *Galileo* spacecraft was associated with these cloud systems.

One reason that the storms can last so long on Jupiter is that there is no solid surface directly below the clouds to interfere with the flow. The atmosphere has no distinct bottom, just shearing due to greater pressure. Thus, the weather pattern is free to flow in response to the giant planet's spin. A nearby solid surface would dissipate the energy of the storm clouds, as happens to hurricanes that make landfall on Earth.

But how can the biggest storms on Jupiter last for centuries within a constantly changing atmosphere? The large ovals and spots can survive by sucking in and engulfing smaller eddies that pass in their vicinity, like leaves in a whirlpool of water, consuming them and extracting their energy. The little short-lived storms feed their energy into the larger storms, just as a large fish eats smaller ones. And the food chain continues to the very top, with giant white ovals, each half the size of the Earth, occasionally merging together to become one.

The big storms engulf small ones, and the small storms probably pull up their energy from hotter, lower depths. As the moist, internal heat rises in the stormy updrafts, it can sustain the swirling clouds, supplying the energy that drives much of Jovian weather. Terrestrial weather systems can similarly include hot rising air as well as cool downdrafts, but the heat on Jupiter is generated from a completely different source – from deep in the planet and not from sunlight.

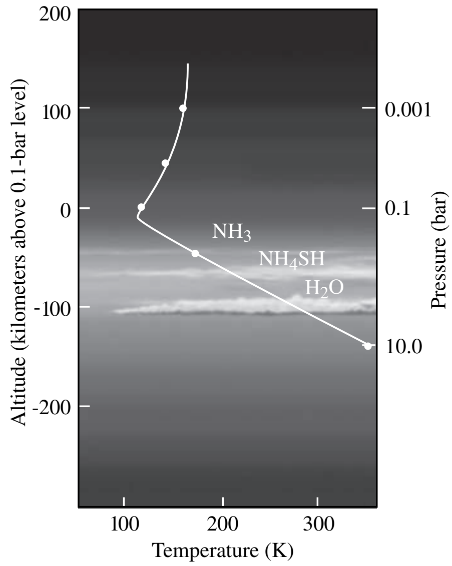


Fig. 9.4 Temperature and pressure at Jovian cloud levels The fading radio signals when the *Voyager 1* and *2* spacecraft passed behind Jupiter in 1979 revealed the temperatures (*bottom axis*) and pressures (*right axis*) in its upper atmosphere. The temperature reaches a minimum of about 114 kelvin at a level called the tropopause where the atmospheric pressure is 0.1 bars, or 100 millibars. By way of comparison, the pressure of the Earth's atmosphere at sea level is 1.0 bar. The altitudes (*left axis*) are relative to the 0.1 bar level, and the dots are spaced to indicate tenfold changes in pressure. Solar radiation causes the temperature to increase with height just above the tropopause. At lower levels, the temperature and pressure increase systematically with depth. Three expected cloud layers of ammonia (NH_3), ammonium hydrosulfide (NH_4SH), and water ice (H_2O) are shown. The altitudes of the predicted cloud layers are based on a gaseous mixture that is of solar composition. An increase of the abundance of a condensable gas by a factor of 3 would lower the altitude of the cloud base by about 10 kilometers.

Cloud layers and colors on Jupiter

Radio signals can penetrate Jupiter's clouds and probe the planet's outer atmosphere, just as radio waves travel from a distant transmitter to your car radio on a cloudy day. Since the weight of overlying layers compresses the gas to greater density at lower depths, the radio signals experience more pronounced refractive alterations when passing through deeper regions of the Jovian atmosphere. These changes have been observed by monitoring homebound radio transmissions from *Voyager 1* and *2* when the spacecraft passed behind the planet, and they have been used to deduce the density, pressure and temperature as a function of altitude in and below the clouds.

If we could descend through Jupiter's thin cloud layer, we would find that the temperature and pressure increase with depth (Fig. 9.4). As in any planetary atmosphere, the atoms and molecules collide more frequently in the increasingly compressed, denser regions of Jupiter's atmosphere, so the pressure and temperature increase there. At the cloud-tops the temperature is a freezing 114 kelvin and the atmospheric pressure is about 0.1 bar, or one-tenth that of the Earth's air pressure at sea level. In slightly deeper layers, about 130 kilometers down, the temperature rises to a balmy 300 kelvin, well above the freezing point of water, at 273 kelvin. In these warmer regions, the pressure is comparable to the air pressure at the surface of the Earth, leading to speculations that living things might reside there (Focus 9.1). And above them it is cold enough to freeze various gases into ice to form the clouds.

Given the profile of temperature and pressure with altitude, it is possible to infer the altitude at which clouds of various types should form. The early calculations, initiated

Focus 9.1 Speculations about life in Jupiter's atmosphere

Once they established the temperatures and pressures in the outer atmosphere of Jupiter, scientists could speculate about the possibility of primitive life existing there. The outer atmosphere may be too cold for life to exist, for it would freeze to death. Deeper down inside the planet it is too hot to even allow organic molecules to exist; they would break apart into their constituent atoms. The molecular constituents of life might nevertheless survive in the region in between, where Earth-like temperatures and pressures exist, perhaps being synthesized from simpler molecules by the action of Jovian lightning.

But the warm part of Jupiter's extended atmosphere contains no solid surface on which primitive creatures could creep or crawl, and strong atmospheric currents would most likely either cycle them up into the frigid heights or drag them down to scalding depths. Heavy organisms would just sink down into the lethal heat. Some imaginative astronomers have therefore argued that buoyant inflated organisms could be floating in Jupiter's outer atmosphere, bobbing up and down like terrestrial jellyfish to seek more clement conditions without sinking too deeply into the planet.

Other researchers have argued that life on Jupiter is very implausible. They reason that biological compounds could not survive the harsh environment. This conclusion was reinforced when the *Galileo* spacecraft dropped a probe into the giant planet, showing that sophisticated organic molecules were not present at the entry site.

Nowadays, many scientists have forgotten about these early speculations about possible life inside Jupiter, and have turned their attention to Jupiter's moon Europa, which might have a life-sustaining ocean beneath its icy surface.

by John S. Lewis (1941–) in 1969, assumed that the gas mixture is in chemical equilibrium and has a uniform composition like that of the Sun. And although there have been many refinements since then, three distinct cloud layers are always predicted. As one proceeds upward from the interior of Jupiter, the temperature and pressure fall to the point where the gases of water, ammonium hydrosulfide and ammonia are expected to condense to form clouds. Water clouds similarly form in the colder, higher parts of the Earth's atmosphere, which has only one layer of cloudy, stormy weather.

The visible cloud tops of Jupiter consist of ammonia ice crystals, which condense out of the atmosphere at the very low temperatures and pressures there (Fig. 9.5). They create graceful white clouds that probably make up the cold, light-colored zones observed from Earth. This is consistent with spectroscopic measurements of abundant ammonia at the cloud tops of Jupiter.

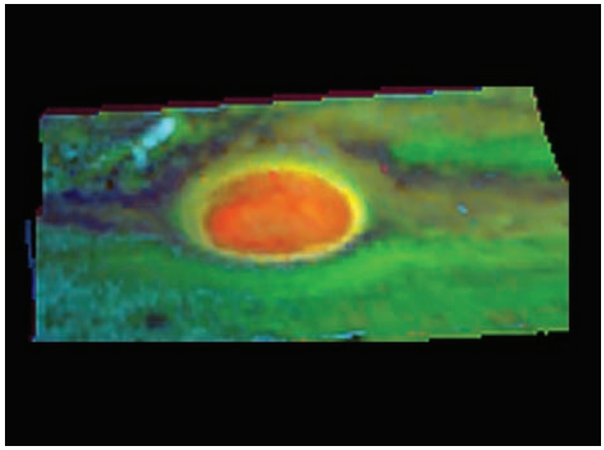


Fig. 9.5 Ammonia ice near Jupiter's Great Red Spot A cloud of ammonia ice (*light blue*) is shown at the northwest (*upper left*) of the Great Red Spot (*middle*), inside its turbulent wake. The cloud was most likely produced by powerful updrafts of ammonia-laden air from deep within Jupiter's atmosphere. Reddish-orange areas show high-level clouds, yellow regions depict mid-level clouds, and green areas correspond to lower-level clouds. Darker areas are cloud-free regions. Light blue depicts regions of middle-to-high altitude ammonia ice clouds. This near-infrared image was taken on 26 June 1996 from the *Galileo* spacecraft. (Courtesy of NASA/JPL.)

Below the ammonia layer, the models predict two other cloud layers. At a depth where the pressure is slightly higher, at about 2 bars, the ammonia combines with hydrogen sulfide to form clouds of ammonium hydrosulfide, which would smell something like rotten eggs if you could get near them. The lowest clouds are predicted to contain water ice, formed at a pressure of about 5 or 6 bars. The water clouds are obscured by the higher cloud cover of ammonia, and are almost never seen from outside the giant planet.

The layered cloud model does not explain Jupiter's highly colorful appearance. Ammonia, ammonium hydrosulfide and water form white ices, so all the expected clouds should be white. Scientists speculate that complex molecules formed by interaction of gases with solar ultraviolet radiation in the outer atmosphere might coalesce and grow to form brown and yellow smog particles. They could account for some of the colors found in the belts – and remind us of a bad smoggy day in Beijing, Los Angeles or Mexico City.

The origin of the red color in Jupiter's clouds is something of a mystery. It might arise when the chemical equilibrium is disturbed by something that energizes

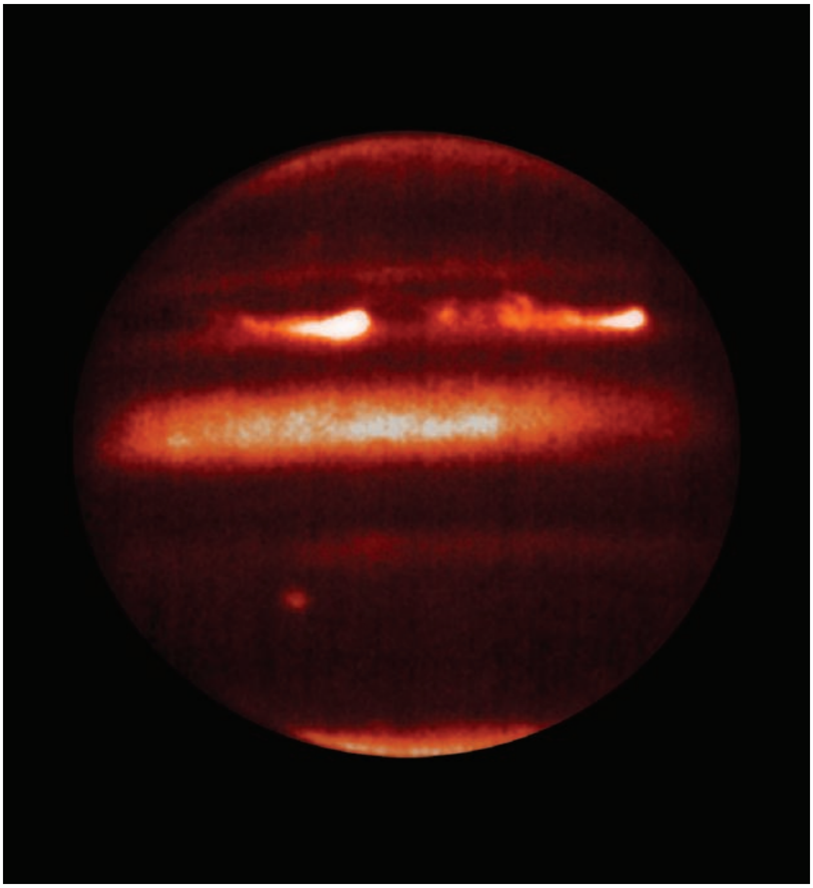


Fig. 9.6 Infrared storms on Jupiter The heat of two continent-sized storms (*top*) is detected in this infrared image. The bright plumes are attributed to storm systems triggered in Jupiter's deep clouds of water ice that move upward in the atmosphere and inject a fresh mixture of ammonia ice and water above the visible cloud tops. Models of the disturbance indicate that the associated jet stream extends deep in the atmosphere, at more than 100 kilometers below the cloud tops, and that the jet stream moves at a speed of about 600 kilometers per hour. This image was obtained from NASA's ground-based InfraRed Telescope Facility (IRTF) in Hawaii on 5 April 2007. (Courtesy of NASA/JPL/IRTF.)

an active chemistry. Lightning bolts, ultraviolet sunlight, high-speed particles, or extreme temperature variations might be responsible. One or more of these sources of energy probably breaks down molecules to produce coloring compounds of sulfur or phosphorus that are present in the atmosphere in only minute quantities. Alternatively, the red colors might be dredged up from greater depths by swirling storms, or perhaps attributed to rare organic, carbon-bearing compounds.

Lightning bolts in wet spots on Jupiter

Ancient mythology was close to the mark when it designated Jupiter as master of the rains, hurling thunderbolts at those who displeased him. Lightning flashes were discovered in *Voyager 1* and *2* images of the dark night-side of Jupiter, apparently illuminating the clouds in massive thunderstorms, and the lightning was confirmed by instruments aboard the *Galileo* spacecraft. Both missions showed that the lightning is concentrated near oppositely directed winds where storm clouds are found.

How deep the lightning occurs can be estimated from its size. The larger the flash, the deeper the lightning discharge. The observed sizes of Jupiter's lightning flashes suggest that they originate from layers in the atmosphere where water clouds are expected to form, at about 100 kilometers down. Only water could condense at these depths. When the *Galileo* cameras followed the night-side lightning sources into the dayside, they confirmed that the lightning originates in deep, moist clouds.

If the lightning bolts on Jupiter are similar to those on Earth, then they probably occur in water clouds where partially frozen water particles become electrically charged. The electrified rain or ice particles rise and fall in the turbulence, causing positive and negative charges to separate. A powerful discharge of current can then flow through the atmosphere, producing the lightning flashes.

So the lightning bolts on Jupiter most likely point to places where there are rapidly falling raindrops and quickly rising air currents. Infrared images of continent-size storms on Jupiter also indicate that they are triggered by upward-moving water clouds (Fig. 9.6). Such a moist convection might transport heat upward, carrying energy into the outer Jovian atmosphere. The circulating air seems to have been detected when *Galileo* sent a probe into the planet's clouds.

Plunging into a dry and windy spot on Jupiter

A pioneering descent into Jupiter's atmosphere took place on 7 December 1995 when an instrument-laden *Galileo Probe* was dropped from the *Galileo* spacecraft into the planet, taking the measure of its composition and winds to well below the visible clouds. The capsule returned data for just over an hour, down to the 20-bar pressure level, until the rising temperatures and crushing pressures wiped the probe out and it disappeared without a trace.

Scientists had expected that the *Galileo Probe* would pass through three cloud layers, composed of different chemicals that condense from tenuous gases at successively higher and colder levels, but contrary to expectations the clouds were not where everyone thought they would be. The probe's instruments saw almost no evidence for clouds. All of the expected cloud constituents were still in the gaseous state, and were found in increasing amounts through and well below the condensation levels where the cloud's ice particles should have been found.

Moreover, the planet was a lot drier than anticipated, at least in the vicinity of the probe-entry site, where the amount of water was five or ten times less than expected. Far less lightning was also detected during the probe's hour-long descent, supporting the conclusion that this part of the upper atmosphere contains little water. The missing clouds and water might be explained if the probe descended into an unusually clear hotspot of dry, downwelling air and reduced cloud cover (Fig. 9.7).

The previous models of Jupiter's clouds were probably too simplified, for they assumed a uniform layering with depth and ignored deep, vertical, up-and-down weather-related activity. Both wet and dry regions are found in the outer atmosphere of Jupiter, just as Earth has tropics and deserts, and they may be related to the circulation of rising and falling air. Winds that rise from the deep atmosphere could dredge up material that lacks water and other cloud-making ingredients. When these winds converge and drop back down, nothing is left to condense back into clouds and a dry clearing is created. Similar downdrafts occur over subtropical deserts on Earth, but, unlike our planet, Jupiter has no firm surface to quickly stop the air's fall.

One part of the weather forecast that proved correct below Jupiter's cloud tops was "windy". Instead of decreasing to a dead calm as the probe descended, the zonal winds stayed strong and even increased with depth. The zonal winds that create the planet's banded appearance continued to whip around the planet at speeds of up to 200 meters per second, until the capsule fell silent and stopped sending readings at about 600 kilometers down. Because little sunlight can penetrate to such depths, the winds must be driven mainly from below. Internal heat, probably left over from planetary formation or contraction of the planet as it slowly cools, is therefore the most likely driving force for Jupiter's powerful winds and ever-changing weather patterns.

Fig. 9.7 Hotspot on Jupiter The dark region near the center of this image is an equatorial "hotspot", similar to the site where the *Galileo* spacecraft parachuted a probe into Jupiter's atmosphere in December 1995. Jupiter has many such regions, and they continually change, so the probe could not be targeted to either hit or avoid them. The dark hotspot is a clear gap in the clouds where infrared radiant energy from the planet's deep atmosphere shines through. Although hotter than the surrounding clouds, these so-called "hotspots" are still colder than the freezing temperature of water. Dry atmospheric gas may be converging and sinking in these regions, maintaining their cloud-free appearance. The bright ovals, shown in other parts of this image, may be examples of upwelling moist air. The images combined in this mosaic were taken on 17 December 1996 from the *Galileo* spacecraft. (Courtesy of NASA/JPL.)

Composition of Jupiter's upper atmosphere

When the *Galileo* spacecraft sent a probe into Jupiter, the relative amounts of several elements were measured in the planet's outermost atmosphere. The abundance of these ingredients has been compared to that of hydrogen, by far the most abundant element in Jupiter, and to the relative amounts found in the outer layers of the Sun (Table 9.2).

Helium, the second most abundant element in both Jupiter and the Sun, was found just a bit depleted from solar amounts, as had been suggested by previous measurements from the *Voyager 1* and *2* spacecraft. The *Galileo* results were more accurate, indicating somewhat more helium than the previous missions had, but still less than the outer layers of the Sun. Helium is apparently removed from Jupiter's outermost atmosphere, by helium raining into the interior of the planet. This slow helium-removal process operates on an awesome scale in neighboring Saturn, whose outer atmosphere is severely depleted in helium.

Heavier elements, such as carbon, nitrogen and sulfur, were enriched in the Jovian atmosphere by a factor of

Table 9.2 Element abundance in the outer layers of Jupiter and the Sun[a]

Element	Symbol	Chemical form	Jupiter	Sun	Jupiter/Sun
Helium	He	Helium	0.078	0.097	0.804
Carbon	C	Methane, CH_4	1.0×10^{-3}	3.6×10^{-4}	2.78
Nitrogen	N	Ammonia, NH_3	4.0×10^{-4}	1.1×10^{-4}	3.64
Oxygen	O	Water, H_2O	3.0×10^{-4}	8.5×10^{-4}	3.53
Sulfur	S	Hydrogen sulfide, H_2S	4.0×10^{-5}	1.6×10^{-5}	2.50
Deuterium	D	Deuterium	3.0×10^{-5}	3.0×10^{-5}	1.0
Neon	Ne	Neon	1.1×10^{-5}	1.1×10^{-4}	0.10
Argon	Ar	Argon	7.5×10^{-6}	3.0×10^{-6}	2.50
Krypton	Kr	Krypton	2.5×10^{-9}	9.2×10^{-10}	2.72
Xenon	Xe	Xenon	1.1×10^{-10}	4.4×10^{-11}	2.50

[a] Number of atoms per atom of hydrogen, designated by the symbol H.

about 3 when compared to a mixture of solar composition. This result had also been anticipated, for the carbon (C) and nitrogen (N) combine chemically with hydrogen (H) to form methane (CH_4) and ammonia (NH_3), and scientists have long known that Jupiter has about three times as much carbon and nitrogen in the form of these gases as the Sun. Still, even with this enrichment, hydrogen and helium comprise 99 percent of Jupiter's substance by volume.

The *Galileo Probe*'s instruments detected surprisingly high concentrations of argon, krypton and xenon. These three chemical elements are called noble gases, perhaps because they are very independent and do not combine with other chemical elements. They are enriched compared with solar composition by about the same factor as carbon, nitrogen and sulfur. In contrast, another noble element, neon, was starkly depleted with about ten times less than the solar amount.

These composition results are related to the formation and subsequent evolution of Jupiter. The apparent enrichment of elements like carbon, relative to light hydrogen, probably occurred when some of the light gas was blown out of the solar system by the active young Sun. In order to catch and retain the noble gases, Jupiter had to freeze them – which is not possible at the giant planet's current distance from the Sun. The noble gases might have been brought in from colder, more distant regions by comet-like bodies that helped build up young Jupiter, or the entire planet might have originated further away from the Sun and gradually migrated inward to where it is now.

The anomalous depletion of neon was even explained before its discovery. Scientists predicted that the neon would dissolve in helium, which rains down inside the planet and takes the neon with it.

9.3 Beneath Jupiter's clouds

Educated guesses about Jupiter's internal constitution

We cannot see beneath the clouds of Jupiter, but we can use external measurements to constrain its internal properties. As an example, we now know that the giant planet emits its own heat radiation, which means that it is hot inside. Since Jupiter and the Sun originated from similar material at the same time, a good initial assumption is that they have the same ingredients with similar proportions. The planet's low bulk density indicates that it is in fact composed largely of hydrogen and helium, just as the Sun is. Jupiter's oblate shape and rapid rotation also tell us something about the way it is constructed inside. Due to the enormous pressures inside Jupiter, most of the planet's hydrogen is compressed into a liquid metallic form, which has been created in the terrestrial laboratory and helps account for the giant's strong magnetic field. All of these constraints have been pieced together to make a picture of Jupiter's invisible interior.

Jupiter is an incandescent globe

With the advent of ground-based infrared measurements of the planets, pioneered by Frank J. Low (1933–2009) and his colleagues in the 1960s, astronomers were surprised to discover in 1969 that the giant planet is an incandescent globe with its own internal source of heat. This result was confirmed in greater detail with instruments aboard the *Voyager 1* and *2* spacecraft, which determined precisely how much infrared heat radiation was emerging from inside the planet. They showed that Jupiter is radiating

Focus 9.2 Stars that do not quite make it

Planets are supposed to shine only by reflected sunlight, while most stars generate their own radiation by thermonuclear reactions in their exceptionally hot cores. Jupiter has its own internal energy source, so it resembles a star in this respect, but unlike a star Jupiter has a relatively cold outer atmosphere. The planet shines by heat left over from its formation rather than nuclear fusion.

The process of gravitational contraction that warmed the inside of young Jupiter to its present central temperature of about 17 000 kelvin also heated the center of the Sun to about 15 million kelvin. That was hot enough to ignite the thermonuclear reactions that make our star shine. The Sun became much hotter inside because it is more massive, weighing in at 1000 times the mass of Jupiter. So Jupiter could have collapsed to form a star if it had more mass, and some astronomers therefore like to call Jupiter a star that didn't quite make it. In fact, calculations indicate that Jupiter might have become a star if it had been only 80 times more massive than it is now.

There are certain stellar objects, known as brown dwarfs, which also do not have enough mass to trigger nuclear fusion reactions in their core. They can shine faintly for about 100 million years as gravitational energy is converted into heat. Objects more massive than about 13 Jupiter masses, but not massive enough to ignite thermonuclear reactions, are thought to be brown dwarf stars, while those with less mass are classified as giant planets.

1.67 times as much energy as it receives from the Sun, and almost half of the total energy that Jupiter loses must come from its interior. That essentially meant that the planet had to be unexpectedly hot inside.

Jupiter must have been still hotter when it formed, thanks to the energy released during the gravitational collapse that accompanied its growth. When the newborn planet coalesced from a larger primordial cloud, gravitational energy was converted into heat as particles and small bodies fell inward and collided with each other. Such a process ignited the internal fires of the Sun and other stars, but Jupiter was not quite massive enough to become a star (Focus 9.2).

The compression inside Jupiter also excites the gas and leads to radiation that can carry off some of the energy. But so much heat is still left inside Jupiter since its time of formation that it is still cooling off, pumping out about twice as much energy as it absorbs from the Sun. In contrast, a small planet like Earth would have radiated away the heat of its formation long ago. The giant planet's enormous size makes it a much better heat-trap than the terrestrial planets.

The Earth is now heated internally by the decay of radioactive elements contained in its rocks, but this has nothing to do with Jupiter's excess heat radiation. The giant planet contains relatively little rocky material, and the observed heat flow is 100 times too large to be explainable by radioactive heat production, even if Jupiter were composed almost entirely of terrestrial rocks rather than mostly hydrogen and helium.

Jupiter could also be shrinking slightly today, converting gravitational energy into heat and supplementing the leftover heat of its initial formation. But we cannot determine if that is happening. The giant planet needs to contract by only one meter per century to supply the observed amount of internal energy radiated by the planet, and astronomers could never measure its radius with that kind of precision.

Ingredients of Jupiter at formation

According to the widely accepted nebular hypothesis, the Sun and planets formed together during the collapse of a rotating interstellar cloud called the solar nebula. Most of it fell into the center, until it became hot enough to ignite the Sun's nuclear fires. Further out, the planets formed out of a whirling disk of the same material.

If the nebular hypothesis is correct, and the whole solar system originated at the same time, then you might expect Jupiter to have a similar chemical composition to the Sun. To a first approximation, the abundance of the elements in the giant planet does indeed mimic that of the Sun, with a predominance of the lightest element hydrogen. It is the most abundant element in most stars, in interstellar space, and in the entire Universe. The second most abundant element in both Jupiter and the Sun is helium, and hydrogen and helium together account for the low bulk density of both objects, at 1326 and 1409 kilograms per cubic meter respectively.

Spectroscopic observations indicate that the outer atmosphere of both Jupiter and the Sun also contain heavier elements, like carbon, nitrogen and oxygen. The giant planet has a bit more than the star, but even with this enrichment, heavy elements comprise less than 1 percent of the planet's composition by volume – all the rest is hydrogen and helium.

At the frigid temperatures where Jupiter and the other giant planets originated, the carbon (C), nitrogen (N), and

Table 9.3 Oblateness of the giant planets and the Earth[a]

Planet	Equatorial radius, R_e (km)	Polar radius, R_p (km)	Oblateness $(R_e - R_p)/R_e$
Earth	6 378.140	6 356.755	0.003 353
Jupiter	71 492	66 854	0.0649
Saturn	60 268	54 364	0.0980
Uranus	25 559	24 973	0.0229
Neptune	24 766	24 342	0.0171

[a] The radii are given in units of kilometers (km). The radii of the giant planets are those at the level where the atmospheric pressure is equal to one bar, the pressure of air at sea level on Earth.

oxygen (O) atoms would have bonded with the abundant hydrogen (H) to form methane (CH_4), ammonia (NH_3) and water (H_2O), respectively. These compounds are known as "ices" because they would have condensed into solid ice at the freezing temperatures far from the Sun. Rocky material, containing atoms of silicon and iron, would also have been present in lesser amounts. The ice and rock are now located at the center of Jupiter, probably because they coalesced to form a relatively massive nucleus that pulled in the surrounding hydrogen and helium, or perhaps because they settled gravitationally into the central core after the planet formed.

At the higher temperatures closer to the Sun, where the Earth and other terrestrial planets formed, the icy material would be vaporized and could not condense. That left only rocky substances to coalesce and merge together to form the terrestrial planets. Their modest mass and proximity to the Sun would not allow them to capture and retain the abundant lighter gases, hydrogen and helium, directly from the solar nebula.

An equatorial bulge

Observations with even a small telescope show that Jupiter is not a sphere. It has a perceptible bulge around its equatorial middle and is flattened at the poles. This elongated oblate shape is caused by Jupiter's rapid spin. The outward force of rotation opposes the inward gravitational force, and this reduces the pull of gravity in the direction of spin. Since this effect is most pronounced at the equator, and least at the poles, Jupiter has an oblate shape that is elongated along the equator. The same thing happens to all the giant planets, and even to the solid Earth (Table 9.3).

Table 9.4 Range of pressures

Location	Relative pressure
Beneath the foot of a water-strider	0.000 01
Inside a light bulb	0.01
Earth's atmosphere at sea level	1.0
Inside a fully charged scuba tank	100.0
Deepest ocean trench	1 000.0
Pressure at which hydrogen becomes metallic	1 000 000.0
Center of Jupiter	70 000 000.0

The amount of Jupiter's non-spherical extension depends on both its rate of rotation and the internal distribution of its material. The faster the spin, the more the outward push and the greater the elongation. And given its rotation, with the rapid period of 9.9249 hours, the size of the equatorial bulge depends on how Jupiter's mass is distributed inside. The more massive the planet's dense core, the smaller the equatorial bulge.

Scientists have measured the properties of Jupiter's equatorial bulge by accurately determining the motion of Jupiter's natural satellites, and closely tracking the *Pioneer 11* and *Voyager 1* and *2* spacecraft as they flew close to the giant. If Jupiter were a perfectly spherical planet, it would act as if all its mass was concentrated in a single central point and the motions of natural satellites or spacecraft would not depend on their orientation with respect to the planet's equator. In contrast, an oblate planet produces an extra force that tugs the moving object toward its equatorial bulge, and also toward any internal core. When combined with the known mass, volume and rotation rate of Jupiter, observations of these effects indicate that Jupiter has a dense core containing up to 12 Earth masses. Such a central object, presumably composed of non-gaseous rock and ice, was apparently required to initiate the accumulation of the giant planet's extensive hydrogen shell, which now compresses the core to high temperatures and pressures.

Enormous pressures and strange matter

The pressure inside Jupiter grows with depth, just as the pressure increases when you dive into the depths of an ocean. The particles at greater depths are compressed into smaller volumes by the weight of overlying material, so they collide with each other more frequently. This causes an increase in pressure at deeper levels inside the giant planet (Table 9.4).

Near Jupiter's cloud tops, the pressure is about the same as the air pressure at sea level on Earth, designated as 1 bar or 10^5 pascal, and the pressure increases to an

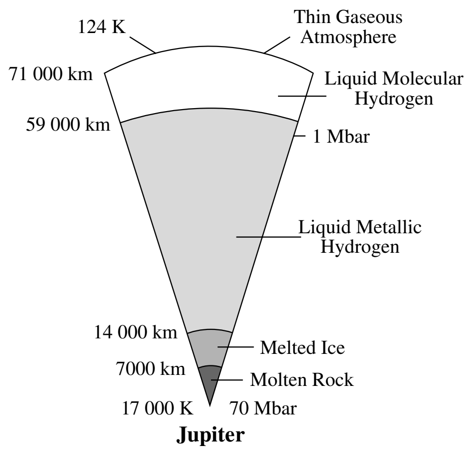


Fig. 9.8 Inside Jupiter Giant Jupiter has a thin gaseous atmosphere covering a vast global ocean of liquid hydrogen. At the enormous pressures within Jupiter's interior, the abundant hydrogen is compressed into an outer shell of liquid molecular hydrogen and an inner shell of fluid metallic hydrogen. The giant planet probably has a relatively small core of melted ice and rock.

astonishing 70 million times that amount at the center of Jupiter. Higher pressures are associated with hotter temperatures, so the temperature also increases with depth inside Jupiter. Out at the one-bar level, the temperature is a freezing 165 kelvin. Deep down at the center, the temperature has risen to 17 000 kelvin, just over three times as hot as the visible disk of the Sun, at 5280 kelvin.

To understand the internal constitution of Jupiter, we need to know what happens to its most abundant ingredient, hydrogen, as the pressure increases. At the low pressure in the outer, visible parts of Jupiter, hydrogen forms a molecular gas, but this atmosphere is just a thin veneer. In proportion, the outer layer of gaseous hydrogen molecules resembles the skin of an apple. Deeper down, where the pressures and temperatures are higher, the hydrogen is liquefied. Indeed the planet is almost entirely liquid (Fig. 9.8). It is mostly just a vast, global drop of liquid hydrogen.

As suggested by Rupert Wildt (1905–1976) in 1938, the intense pressures and temperatures deep inside Jupiter will cause hydrogen molecules to break down, forming something he called "metallic hydrogen". Jupiter is indeed so massive that most of the fluid hydrogen inside is believed to be squeezed into metallic form. Below about one-seventh of Jupiter's radius down, the internal pressure exceeds 1 million bars (1 Mbar) and the liquid molecular hydrogen is transformed into liquid metallic hydrogen. It is said to be in a metallic state because, like a metal, it is an excellent conductor of electricity.

The hydrogen molecules are pushed so closely together that the electrons are squeezed free of any single atom or molecule. These mobile electrons can travel freely from one place to another, moving about and conducting electricity like the electrons in metals such as copper. But unlike a hard, shiny metal, the metallic hydrogen inside Jupiter has been melted at the high temperatures down there, and the molten metal behaves like a liquid rather than a solid.

Most of Jupiter is in the form of liquid metallic hydrogen. And underneath it all is a relatively small core of molten rock and ice. This relatively little core is up to 12 times heavier than the Earth.

The vast liquid shell of metallic hydrogen, which lies just outside Jupiter's core, is no longer just a theoretical conjecture. By re-creating extreme conditions similar to those inside Jupiter, modern laboratory experiments have turned liquid hydrogen into a metal. They have succeeded in pressurizing liquid hydrogen to 1.4 million bars and 3000 kelvin, squeezing the hydrogen into a liquid metallic state that conducts electricity just like any solid metal.

Electrical currents, driven by Jupiter's fast rotation within its liquid metallic shell, apparently generate the planet's strong magnetic field, in much the same way that currents in the Earth's molten metallic outer core produce our planet's magnetism. Jupiter's magnetic field is much more powerful than Earth's magnetism, with a magnetic moment that is 20 000 times as large and a cloud-top strength that is about 14 times Earth's surface magnetic field strength. The greater strength of Jupiter's magnetism could be attributed to the planet's faster rotation, its more extensive metallic region, and the relative proximity of the internal electrical currents to the cloud tops. By way of comparison, Earth's magnetic field is produced within a much smaller metallic core, which extends only halfway to the surface.

9.4 Introduction to the Galilean satellites

Galileo Galilei (1564–1643) discovered Jupiter's four largest moons in January 1610, using the newly invented telescope. They are bright enough to be seen in a pair of binoculars or a small telescope, and if it were not for the glare of Jupiter, these moons would be visible to the unaided eye.

These objects are now collectively called the Galilean satellites, even though Galileo wanted to name them after the Medici family of Firenze. They retain the individual names given to them by Simon Marius (1573–1624), also in 1610. In order of increasing distance from the giant planet, they are Io, Europa, Ganymede, and Callisto, all the names of mythological consorts of Zeus, the Greek equivalent of the Roman god Jupiter.

Table 9.5 Properties of the Galilean satellites[a]

Satellite	Distance from Jupiter center (Jovian radii)	Orbital period[b] (days)	Mean radius (km)	Mass (10^{22} kg)	Mean mass density (kg m^{-3})
Io	5.95 R_J	1.769	1821.61	8.930	3528
Europa	9.47 R_J	3.551	1560.8	4.799	3013
Ganymede	15.1 R_J	7.155	2631.2	14.815	1942
Callisto	26.6 R_J	16.69	2410.3	10.757	1834

[a] The mean distances from the center of Jupiter are in units of Jupiter's equatorial radius, $R_J = 71\,492$ kilometers. The radii are given in units of kilometers (km), the mass is given in kilograms (kg) and the mass density is in units of kilograms per cubic meter (kg/m^3). By way of comparison, the radius of our Moon is 1738 kilometers and the Moon's mean mass density is 3344 kg/m^3. The planet Mercury has a radius of 2439 kilometers and a mean density of 5430 kg/m^3.
[b] The orbital period of Europa is about twice that of Io, and the orbital period of Ganymede is nearly twice that of Europa.

Zeus changed the mortal Io, a beautiful river nymph, into a cow to hide her from his jealous wife. The Ionian Sea is named after the sea that Io the cow swam in during her wanderings. Europa, a Phoenician princess, bore Jupiter three sons, including Minos, the legendary ancestor of the Minoan civilization. Charlemagne subsequently named the continent which he had conquered "Europe", after the young lady. Ganymede was a beautiful Trojan boy, carried off by an eagle to be cupbearer of the gods. The nymph Callisto also conceived one of Zeus's sons, but Zeus's enraged wife turned her into a bear. Callisto and her son were placed in the heavens as the constellations Ursa Major and Ursa Minor, the big and little bears. Parts of these constellations are also known as the Big Dipper and the Little Dipper.

Many of the surface features on the Galilean satellites are named after persons or places in worldwide mythology, including those in the myths of Io, Europa, Ganymede and Callisto. Gods of fire and volcanoes from many cultures were also used for Io – this was changed after scientists realized the extent of the volcanism.

The Galilean satellites provided the first clear example of objects moving about a center other than the Earth, and for this reason they played an important role in the eventual acceptance of Copernicus' model of the solar system. They move in nearly circular orbits near Jupiter's equatorial plane with periods of days, revolving around the planet so quickly that their positions can be seen to change from hour to hour.

The first quantitative physical studies of these worlds became possible during the 19th century when Pierre Simon de Laplace (1749–1827) derived the satellite masses from their mutual gravitational perturbations. When large ground-based telescopes were constructed, astronomers could measure the sizes of these moons, and make approximate estimates of their mean mass densities accurate to roughly 10 percent. Precise determinations of these physical parameters became possible as the result of the *Voyager 1* and *2* flybys of Jupiter in 1979 and the *Galileo* orbital mission from 1995 to 2003 (Table 9.5).

When the two *Voyager* spacecraft flew past Jupiter in 1979, they got only a brief look at the Galilean satellites. However, it was time enough for their cameras to turn the four moons into astonishing places, including active volcanoes on Io, smooth ice plains on Europa, grooved terrain on Ganymede, and the crater-pocked surface of Callisto (Fig. 9.9). The incredible complexity and rich diversity of their surfaces, which rival those of the terrestrial planets, are only visible by close-up scrutiny from nearby spacecraft. Ground-based telescopes provide only a blurred view of the tiny, distant moons.

The volcano-ravaged surface of Io is being transformed before our very eyes, as spacecraft catch volcanoes in the act of erupting and watch lava flowing across its surface. Io must be at least partly molten inside to account for this rampant volcanism. The remarkably smooth surface of Europa, which is nearly devoid of large craters, also suggests a warm, active interior for this satellite. It has emitted material onto the surface that has erased crater-forming impacts in the recent past. An ocean of liquid water, or at least slushy ice, apparently exists at shallow depths beneath Europa's frozen surface, lubricating overlying ice rafts and oozing out into cracks in its icy covering.

The smaller and innermost of the Galilean satellites, Io and Europa, have the highest mass density, which is comparable to that of the rocks found on Earth. These

Fig. 9.9 The Galilean satellites A composite of the four largest moons of Jupiter, which are known as the Galilean satellites. From left to right and increasing distance from Jupiter, the moons shown are Io, Europa, Ganymede, and Callisto. The images of Ganymede and Io are from the *Galileo* spacecraft, while those of Callisto and Europa are from the *Voyager 1* or *2* spacecraft. Io is subject to the strongest tidal stresses from the massive planet. These stresses generate internal heating, which is released at the surface and makes Io the most volcanically active body in our solar system. Europa appears to be strongly differentiated inside with a rock/iron core, a surface layer of bright water ice, and possibly local or global zones of liquid water between these layers. Tectonic resurfacing brightens terrain on the less active and partially differentiated moon Ganymede. Callisto, furthest from Jupiter, appears heavily cratered and shows no evidence of internal activity. (Courtesy of NASA/JPL/DLP.)

satellites are about the same size as our Moon, and have about the same mean mass density. In comparison, the larger, outer satellites Ganymede and Callisto are about the size of Mercury, but much less dense. The low mean mass densities indicate that they are composed in part of water ice. A mean mass density of 2000 kilograms per cubic meter would, for example, be explained if an object consists of half silicate rock and half water ice, with respective mass densities of 3000 and 1000 kilograms per cubic meter.

The compositions of the Galilean satellites were most likely affected by their relative proximity to Jupiter, in much the same way that the ingredients of the planets are related to their distances from the Sun. In both instances, the relatively small, dense objects are close to the center and larger, less dense objects are found farther out. It was probably too warm near the newborn Jupiter for water to condense, explaining why Io and Europa are largely composed of rock. Europa is now covered by ice, which may overlay an ocean of liquid water, but this blanket may be a relatively thin veneer. Recent models suggest that the Galilean moons grew gradually as material was added to a rotating disk surrounding young Jupiter, keeping the temperature low enough for ice in the region of Ganymede and Callisto.

In contrast, the relative cold of regions farther from Jupiter permitted Ganymede and Callisto to retain their ice and become mixtures of ice and rock. This would explain their low mass density, as well as the fact that they are more massive than Io and Europa.

The high reflectivity of the Galilean satellites, combined with the very cold temperatures at their remote distances from the Sun, has long suggested that ice might be present on their surfaces. In fact, Europa is so bright, and reflects so much incident sunlight, that it ought to be covered by pure water ice. Spectroscopic observations in the 1970s, using Earth-based telescopes at infrared wavelengths, indeed identified the expected water ice on the surfaces of Europa, Ganymede and Callisto. Although Io has the high reflectivity one might expect from an ice-covered sphere, the infrared observations failed to detect any signs of water ice. Instead, sulfur dioxide frost was identified on Io's surface by its spectroscopic signature.

Between 1995 and 2003, the orbiting *Galileo* spacecraft sharpened our view of Jupiter and its largest moons, providing captivating images of their diverse surfaces. The spacecraft also effectively looked beneath the surfaces of the four largest moons, inferring their interior structure from their gravitational forces and magnetic fields. It flew close enough to each satellite to measure small changes in the spacecraft's trajectory produced by each moon's gravitational forces, and this provided information about how the satellite's mass is distributed inside (Fig. 9.10). Io, Europa and Ganymede all have a large metallic core, a rocky silicate mantle, and an outer layer of either water ice, for Europa and Ganymede, or rock, for Io. In contrast, Callisto is a relatively uniform mixture of ice and rock.

Additional evidence for a metallic core inside Ganymede was provided by the *Galileo* spacecraft's

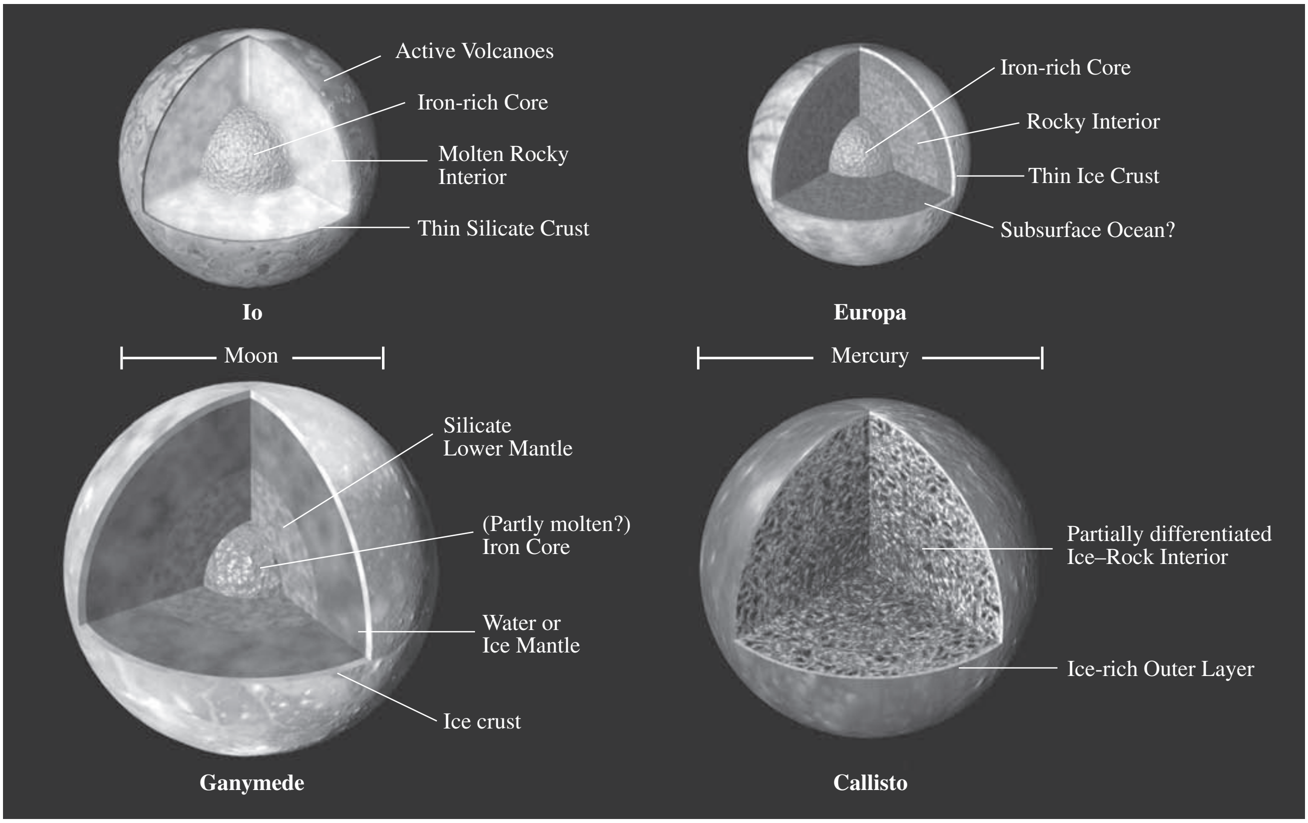


Fig. 9.10 Inside the Galilean satellites Cutaway views of the possible internal structures of the Galilean satellites. Ganymede is at the lower left, Callisto at the lower right, Io on the upper left, and Europa on the upper right. The surfaces of the satellites are mosaics of images obtained in 1979 by the *Voyager 1* or 2 spacecraft, and the interior characteristics are inferred from gravity field and magnetic field measurements from the *Galileo* spacecraft. With the exception of Callisto, all the satellites have iron-nickel cores surrounded by rock shells. Io's rock or silicate shell extends to the surface, while the rock layers of Ganymede and Europa are surrounded by shells of water ice or liquid water. Callisto seems to be a relatively uniform mixture of comparable amounts of ice and rock. *Galileo* images of Europa suggest that a liquid water ocean might now underlie a surface layer of water ice. (Courtesy of NASA/JPL.)

discovery of its intrinsic magnetic field, which ought to be generated within a massive molten core. The metallic cores of Io, Europa and Ganymede probably originated when the satellites were molten inside, the heavier material sinking toward the center and the lighter material rising toward the surface.

9.5 Jupiter's volcanic moon Io

Io: a world turned inside out

The innermost Galilean satellite, Io, has a radius and density that are nearly identical to those of our Moon, but contrary to expectation there are no impact craters on Io. The dramatic landscape is instead richly colored by hot, flowing lava and littered with the deposits of volcanic eruptions (Fig. 9.11). The active volcanoes emit a steady flow of lava that fills in and erases impact craters so fast that not a single one is left.

Io's volcanoes are literally turning the satellite inside out. Each volcano can churn out 100 cubic meters of lava every second, fast enough to fill an Olympic-sized swimming pool every minute, and the active volcanoes collectively provide 45 000 tons (45 million kilograms) of lava every second. They eject enough material to cover the satellite's surface to a depth of 100 meters in a short span of a million years or less. So all of the material that we now see on Io was probably deposited there less than a million years ago. Evidently Io's mantle and crust have been recycled many times over the span of Io's history.

Sulfur and sulfur dioxide give rise to Io's colorful appearance. Its red and yellow hues are attributed to different forms of sulfur, probably formed at different temperatures. Volcanic plumes of sulfur dioxide gas fall and freeze onto the surface, forming white deposits that were first detected by ground-based infrared spectroscopy in the 1970s.

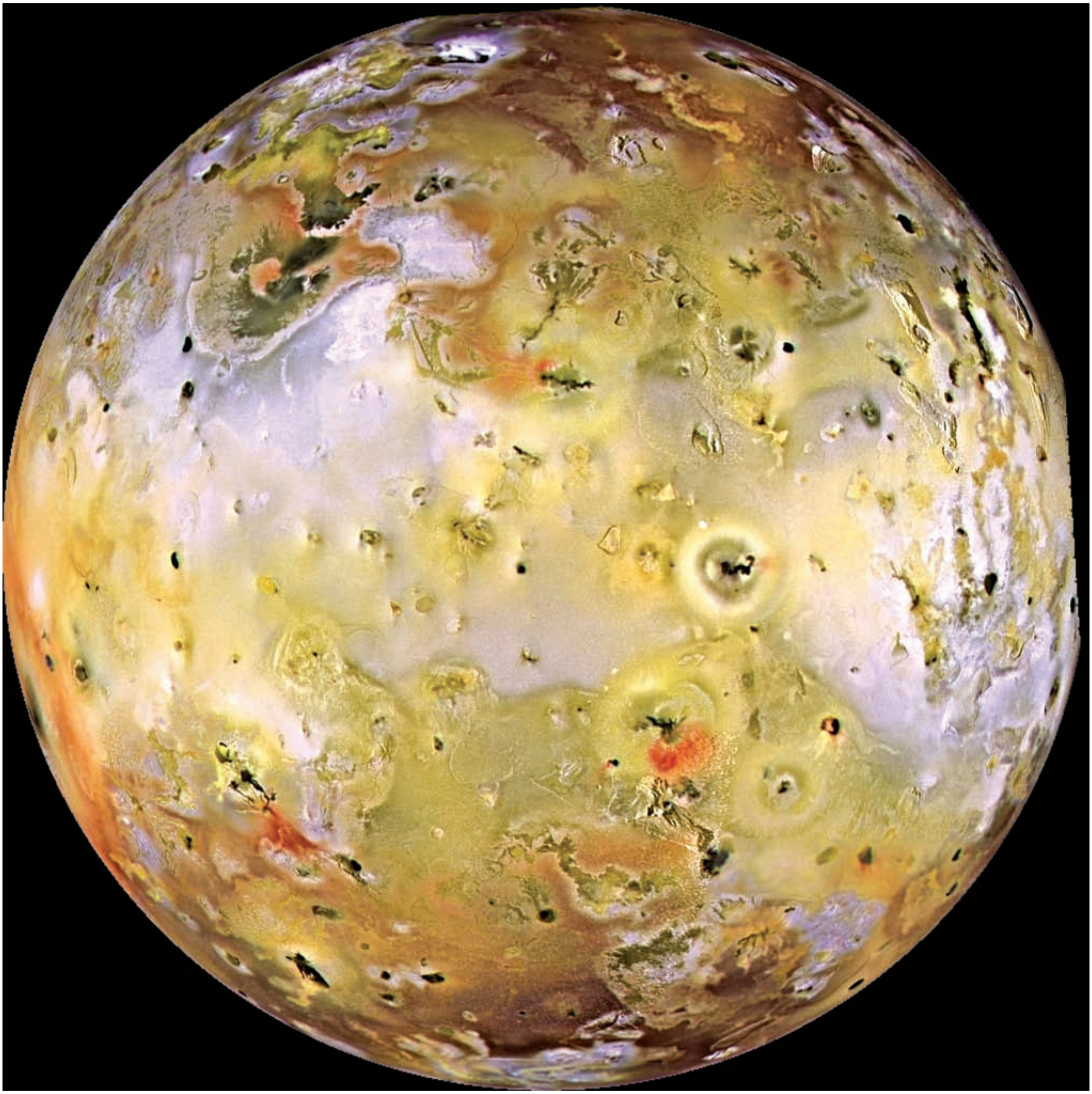


Fig. 9.11 Jupiter's volcanically active moon Io This composite mosaic shows Jupiter's moon Io in true colors, at approximately what the human eye would see. Black, brown, green, orange and red spots mark active volcanic centers. These volcanoes are turning the satellite inside out, continuously forming and re-forming its surface by lava flows. The absence of impact craters suggests that the entire surface is covered with new volcanic deposits much more rapidly than craters are created. This image was taken from the *Galileo* spacecraft on 3 July 1999. (Courtesy of NASA/JPL/U. Arizona.)

Whereas our Moon has been geologically inactive for eons, Io is the most volcanically active body in the solar system (Fig. 9.12). It exhibits gigantic lava flows, fuming lava lakes, and high-temperature eruptions that make Dante's *Inferno* seem like another day in paradise. Scientists estimate that Io has about 300 active volcanoes, and the hotspots of at least 100 of them have been observed.

The cameras aboard the *Voyager 1* spacecraft discovered nine active volcanoes during its flyby in 1979, and the most active volcanoes, such as Prometheus, Loki and Pele, were observed from the *Galileo* spacecraft two decades later (Fig. 9.13). Prometheus is the "Old Faithful" of Io's many volcanoes, remaining active every time it has been observed. Loki is the most powerful volcano in the solar

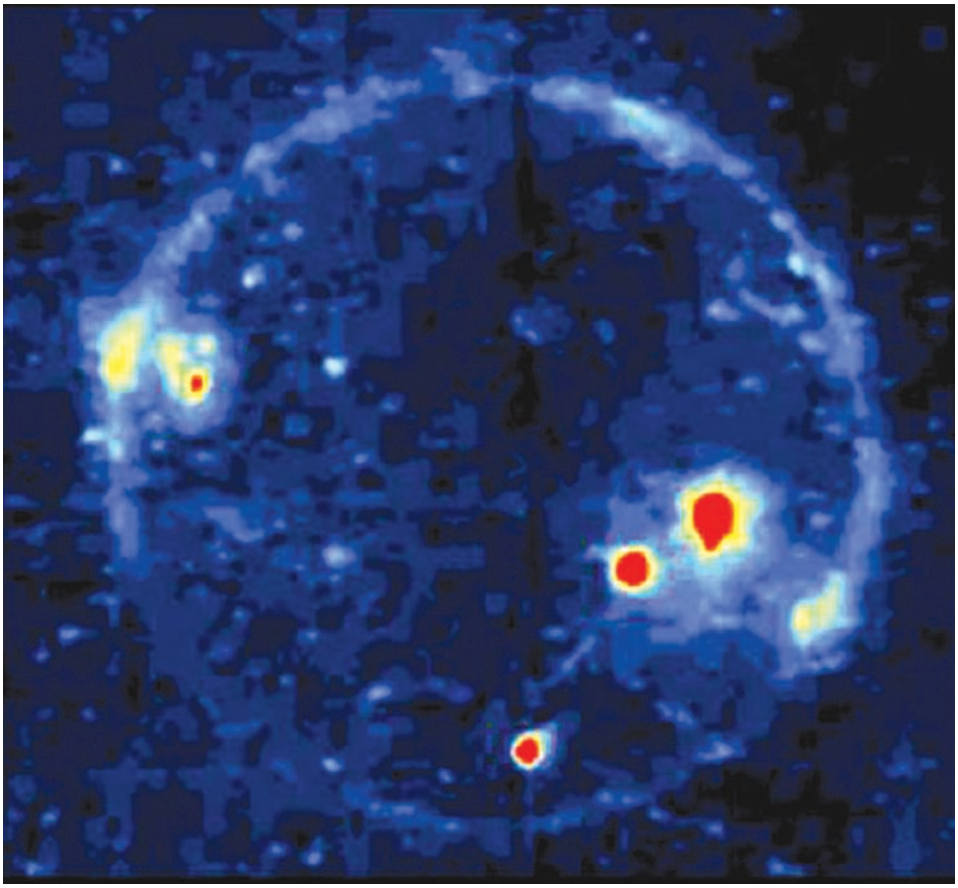


Fig. 9.12 Io's volcanoes glow in the dark This view of Io was taken when the moon was in Jupiter's shadow. The image is color coded so blue to yellow to red represents increasing brightness. The bright spots indicate the locations of volcanic vents on Io, which are spewing hot lava. The brightest spot (*right*) is from the volcanic caldera Pillan Patera with a temperature that exceeds 1700 kelvin, at least 200 kelvin hotter than the temperatures of the hottest volcanic eruptions on Earth today. The volcano Pele is just below and to the left of Pillan, Acaia is at the far left, and Svarog is near the bottom. This image was taken from the *Galileo* spacecraft in 1998. (Courtesy of NASA/JPL/U. Arizona.)

system, consistently putting out more heat than all of Earth's active volcanoes combined. And Pele (Fig. 9.14), the first volcano to be seen in eruption on Io, has repeated the performance for *Galileo* and the *Hubble Space Telescope*. More than 100 active volcanoes have been discovered on Io; many of them have been named for gods of fire, the Sun, thunder and lightning (Table 9.6).

Plumes of volcanic gas erupt from Io's active volcanic vents, rising up to 500 kilometers above the surface. They spread out in graceful, fountain-like trajectories, depositing circular rings of material about 1000 kilometers in diameter. Instruments aboard *Galileo* have practically smelled the hot, sulfurous breath of the eruptions, monitoring the sulfur dioxide gas as it rises, cools and falls. Diatomic sulfur, consisting of two sulfur atoms joined in pairs, has also been detected gushing out of the active volcanoes by instruments on the *Hubble Space Telescope*.

Although some of the kaleidoscopic colors on Io's surface are attributed to sulfur, the bulk of the lava is melted silicate rock. Instruments on the ground and in space have taken the temperature of the searing lava, showing that it sizzles at temperatures of 1700 to 2000 kelvin. That is hotter than any surface temperature of any planetary body in the solar system, even Venus at 735 kelvin. These temperatures rule out substances that melt at lower temperatures, such as liquid sulfur.

So the volcanic vents on Io must be spewing out melted rock, somewhat like terrestrial volcanoes whose lava is rich in iron, magnesium and calcium silicates, but much hotter than lava from Earth's volcanoes. The ubiquitous high-temperature volcanism on Io also has nothing to do with water, a common propulsive agent for some terrestrial volcanoes. There is no evidence that any water now exists on Io, and there may never have been any there. Io is close enough to Jupiter that heat received from the young planet during the satellite's formation may have kept water from condensing. And if any water managed to collect in Io, the interior heat would have probably boiled it away.

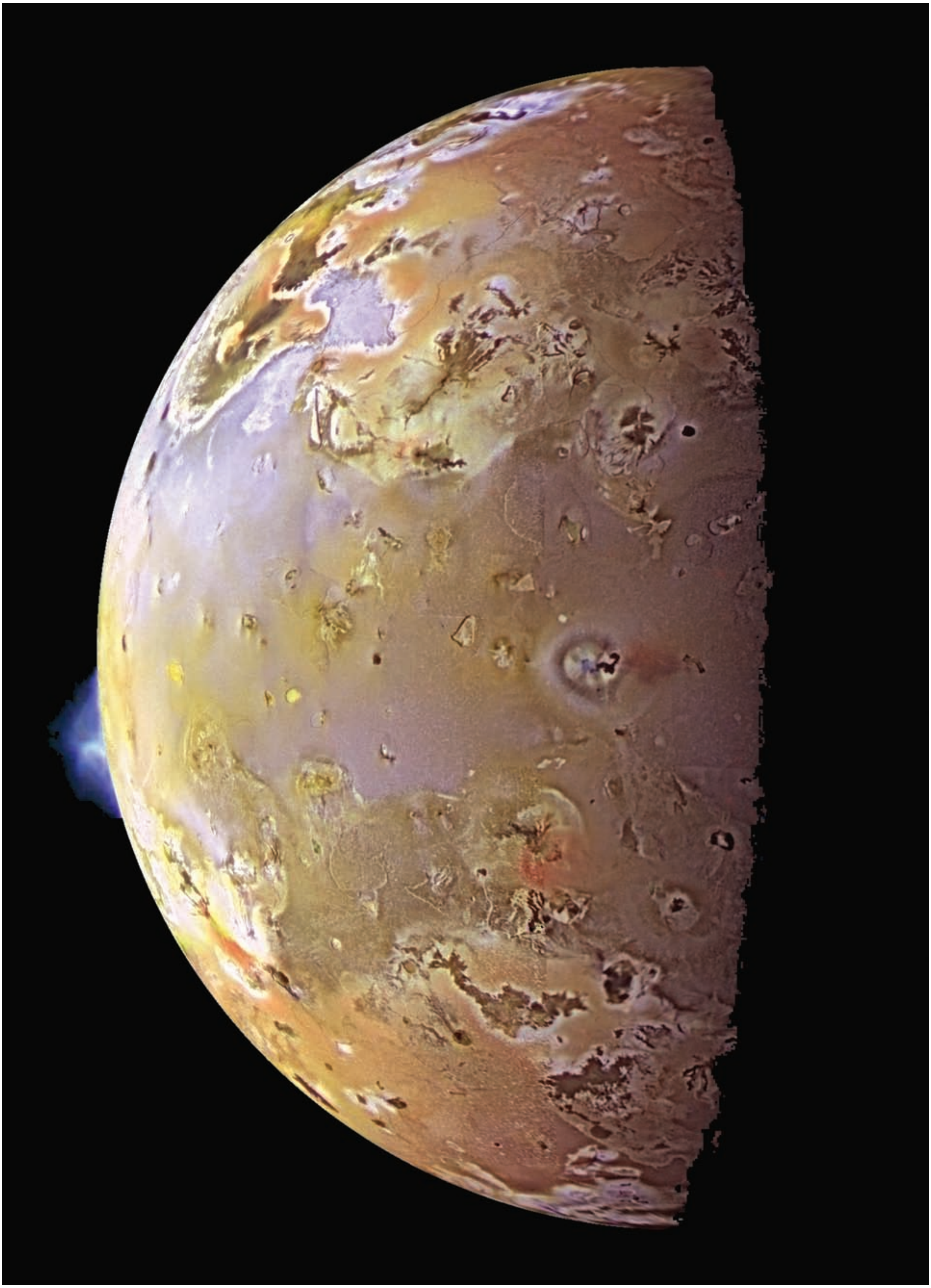


Fig. 9.13 Volcanic eruptions on Jupiter's moon Io A bluish plume rises about 140 kilometers over the bright limb, or edge, of Io (*left*), above the volcanic depression, or caldera, named Pillan Patera. In the middle of the image, near the night/day shadow line or terminator (*right*), the ring-shaped Prometheus plume is seen rising 72 kilometers above Io while casting a shadow to the right of the volcanic vent. Named after the Greek god who gave mortals fire, the Prometheus plume is visible in every image of Io from the *Galileo* spacecraft, including this one acquired in 1997, as well as every *Voyager* image of Io acquired 18 years before in 1979, suggesting continued activity from the same volcano for decades. (Courtesy of NASA/JPL/U. Arizona.)

Table 9.6 Major eruptive volcanic centers on Io

Name	Latitude (degrees)	Longitude (degrees)	Origin of name
Amirani	24.5N	114.7W	Georgian god of fire
Kanehekili	18.2S	33.6W	Hawaiian thunder god
Loki	18.2N	302.6W	Norse blacksmith, trickster god
Marduk	29.3S	209.7W	Sumero-Akkadian fire god
Masubi	49.6S	56.2W	Japanese fire god
Maui	19.5N	122.3W	Hawaiian demigod who sought fire
Pele	18.7S	255.3W	Hawaiian goddess of the volcano
Prometheus	1.3S	153.9W	Greek fire god
Surt	45.2N	336.5W	Icelandic volcano god (Surter)
Thor	39.2N	133.1W	Norse god of thunder
Volund	28.6N	172.5W	Germanic supreme smith of the gods
Zamana	18.4N	172.6W	Babylonian Sun, corn, and war god

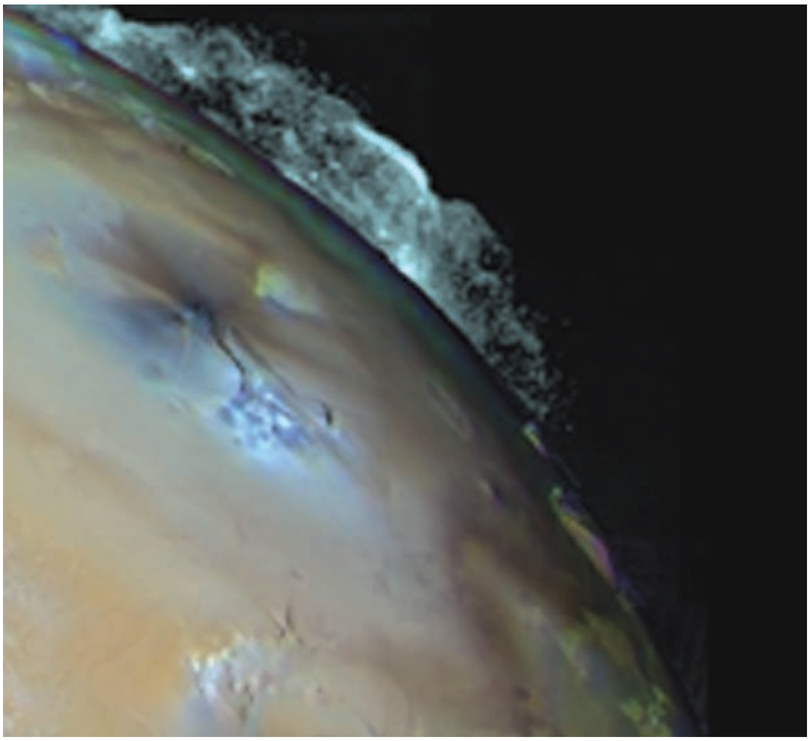


Fig. 9.14 Volcano Pele erupts on Jupiter's moon Io During its flyby on 4–5 March 1979, the *Voyager 1* spacecraft captured this image of an active volcano on Jupiter's energetic moon Io. The volcano has been named Pele, after the Hawaiian goddess of the volcano. Its erupting plume is visible at the upper right, rising to a height of about 300 kilometers above the surface in an umbrella-like shape. The plume has been ejected from the triangular-shaped blue and white complex of hills (*right center*). In this enhanced color image, we see the plume fallout as concentric brown and yellow rings, the largest stretching across 1400 kilometers and covering an area the size of Alaska. Pele remained active for at least two decades, when the *Galileo* spacecraft imaged new deposits from its plumes in the later 1990s. Large tidal distortions raised in Io by Jupiter heat the moon's interior, and the hot magma then expands, rises and forces its way out through volcanoes. (Courtesy of NASA/JPL/USGS.)

Io's tides of rock

What is keeping Io hot inside, warming up its interior, melting its rocks, and energizing its volcanoes? The heat released during the moon's formation and subsequent radioactive heating of its interior should have been lost to space long ago, just as our Moon has lost the internal heat of its youth and become an inert ball of rock. Unlike the Earth, whose volcanoes are energized by heat from radioactivity and friction due to mass motion, it is tidal distortions, created by Jupiter and its other moons, which sustain Io's molten state.

Just as the gravitational force of the Moon pulls on the Earth's oceans, raising tides of water, the gravitational force of massive Jupiter creates tides in the rocks of Io. Since the pull of gravity is greatest on the closest side to Jupiter, and least on the farthest side, Io's solid rocks are drawn into an elongated shape. But this tidal distortion does not melt the rocks by itself. If Io remained in a circular orbit, one side of the moon would always face Jupiter, its tidal bulges would not change in height, and no heat would be generated.

Shortly before the *Voyager* spacecraft encountered Io in 1979, Stanton Peale (1937–), Patrick Cassen (1940–), and Raymond Reynolds noticed that Io's orbit is slightly out of round, and predicted that the resultant tidal flexing of Io would cause "widespread and recurrent volcanism". The three Galilean satellites Io, Europa and Ganymede resonate with each other in a unique orbital dance, known as the Laplace resonance, in which Io moves four times around Jupiter for each time Europa completes two circuits and Ganymede one. This congruence allows small forces to accumulate into larger ones.

Although firmly gripped in Jupiter's gravity Io is also yanked in the opposite direction by the repetitive gravitational pulls of Europa and Ganymede. The resultant gravitational tug-of-war between Jupiter and the satellites distorts the circular orbits of all three moons into more elliptical ones. The effect is greatest for Io, which revolves nearest to Jupiter, but there is a noticeable consequence for Europa and perhaps even Ganymede.

During each lap around its slightly eccentric orbit, Io moves closer to Jupiter and further away, wobbling back and forth slightly as seen from Jupiter (Fig. 9.15). The strong gravitational forces of the planet squeeze and stretch Io rhythmically, as the solid body tides rise and fall. Friction caused by this flexing action heats the material in much the same way that a paper-clip heats up when rapidly bent back and forth. This tidal heating melts Io's interior rocks and produces volcanoes at its surface.

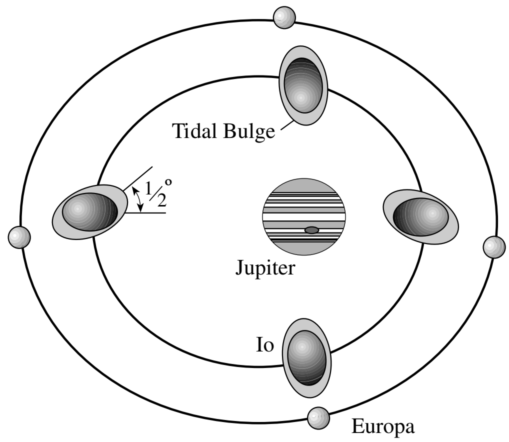


Fig. 9.15 Tidal flexing of Io Due to an orbital resonance with nearby Europa, Jupiter's satellite Io has a non-circular orbit. The forced eccentricity makes Io travel at different speeds along its orbit, and the side facing Jupiter nods back and forth slightly, as seen from the planet. Although only half a degree in extent, this movement causes varying tidal forces inside the satellite, flexing it in and out like squeezing an exercise ball with your hand. This, in turn, generates internal friction and heat, leading to the active volcanoes seen on Io with instruments aboard the *Voyager 1* and *2* and *Galileo* spacecraft. In this drawing, Io's size and the eccentricity of its orbit are exaggerated when compared with Jupiter.

Magnetic connections between Jupiter and Io

Earth-based observations in the 1970s revealed a vast cloud of sodium atoms that envelops Io, forming an extended atmosphere that is nearly as big as Jupiter. The sodium cloud stretches backward and forward along Io's orbit, until the sodium atoms become ionized and no longer emit the light that makes them visible. The neutral, or un-ionized, sodium atoms have probably been chipped off the surface of Io by the persistent hail of high-energy particles found near the giant planet.

The volcanoes on Io provide the raw material for the satellite's tenuous atmosphere of sulfur dioxide (SO_2) that gathers above the erupting vents like localized umbrellas. The volcanic plumes are like fountains, with eruptions that arch gracefully back to Io's surface, and the gas is not propelled with sufficient velocity to escape the satellite's gravitational pull. Nevertheless, atoms of sulfur (S) and oxygen (O) can escape from Io once they are ionized by exposure to radiation from the Sun or from the hail of energetic particles in Io's vicinity. These ions have been detected from the *Voyager* and *Galileo* spacecraft by their ultraviolet glow.

Since charged particles cannot cross magnetic field lines, Jupiter's spinning magnetic field confines and directs the sulfur and oxygen ions into a doughnut-shaped ring known as the plasma torus (Fig. 9.16). As the giant planet rotates, it sweeps its magnetic field past Io, stripping off about a ton, or 1000 kilograms, of sulfur and oxygen ions every second. This material is lost from Io forever, and is continuously replenished by its volcanic activity, albeit indirectly through subsequent ionization.

Once coupled to the Jovian magnetic field, the sulfur and oxygen ions are accelerated to high velocity. Carried by the field, which is anchored inside Jupiter, the ions revolve around the planet once every 9.9249 hours, while Io orbits Jupiter in a more leisurely period of 42.48 hours. So the ions are always catching up with the satellite, and some of them slam into its surface, dislodging and energizing material and lifting it into the thin atmosphere.

As Jupiter's magnetic field sweeps past Io, it generates an enormous electrical potential of 400 000 volts, allowing

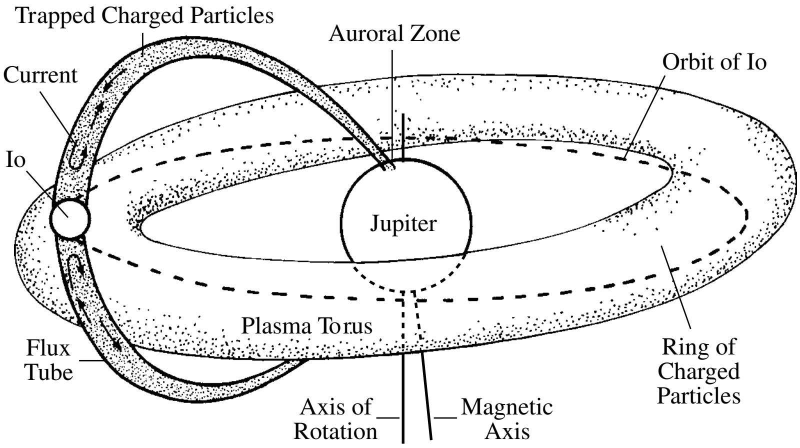


Fig. 9.16 Flux tube and plasma torus An electric current of 5 million amperes flows along Io's flux tube. It connects Io to the upper atmosphere of Jupiter, like a giant umbilical cord. The plasma torus is centered near Io's orbit, and it is about as thick as Jupiter is wide. The torus is filled with energetic sulfur and oxygen ions that have a temperature of about 100 000 kelvin. Because the planet's rotational axis is tilted with respect to the magnetic axis, the orbit of the satellite Io (*dashed line*) is inclined to the plasma torus. Currents are generated as the plasma from the Io torus spreads into the vast, rotating magnetosphere of Jupiter, and these currents couple the moon to Jupiter's atmosphere where they stimulate a ring, or oval, of aurora emissions.

a powerful electric current of 5 million amperes to flow from Io to the poles of Jupiter and back again. The electrons move along Jupiter's magnetic field lines, within a magnetic flux tube that attaches the moon to its planet like a giant electromagnetic umbilical cord. Instruments aboard the *Galileo* spacecraft have detected beams of electrons flowing along the flux tube. They generate an awesome natural power of 2.5 trillion watts, vastly exceeding that of any terrestrial energy-generation plant.

When the electrons in this huge electrical circuit collide with the atoms in Io's tenuous atmosphere, they generate a dazzling light show of red, green and blue emissions (Fig. 9.17). And when the electrons are directed into the atmosphere of Jupiter, at the opposite end of the circuit, they trigger its bright aurora emissions, marking the glowing foot of the flux tube. Currents in this cosmic power-station also generate powerful bursts of radio noise, noticed since 1964, which are strongly controlled by Io's orbital position.

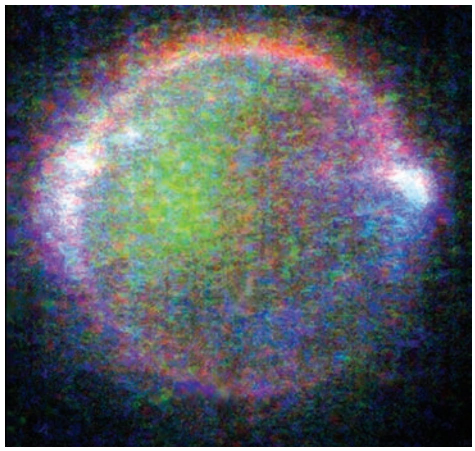


Fig. 9.17 Aurora on Jupiter's moon Io The ghostly glow of aurora are detected in the atmosphere of Io when the satellite is in Jupiter's shadow. The aurora is produced by energetic particles, which are trapped in Jupiter's magnetic field and collide with Io's atmospheric gases, creating a red and green glow analogous to the aurora of Earth. Blue light is caused by dense volcanic plumes and may indicate regions that are electrically connected to Jupiter itself. (Courtesy of NASA/JPL/U. Arizona.)

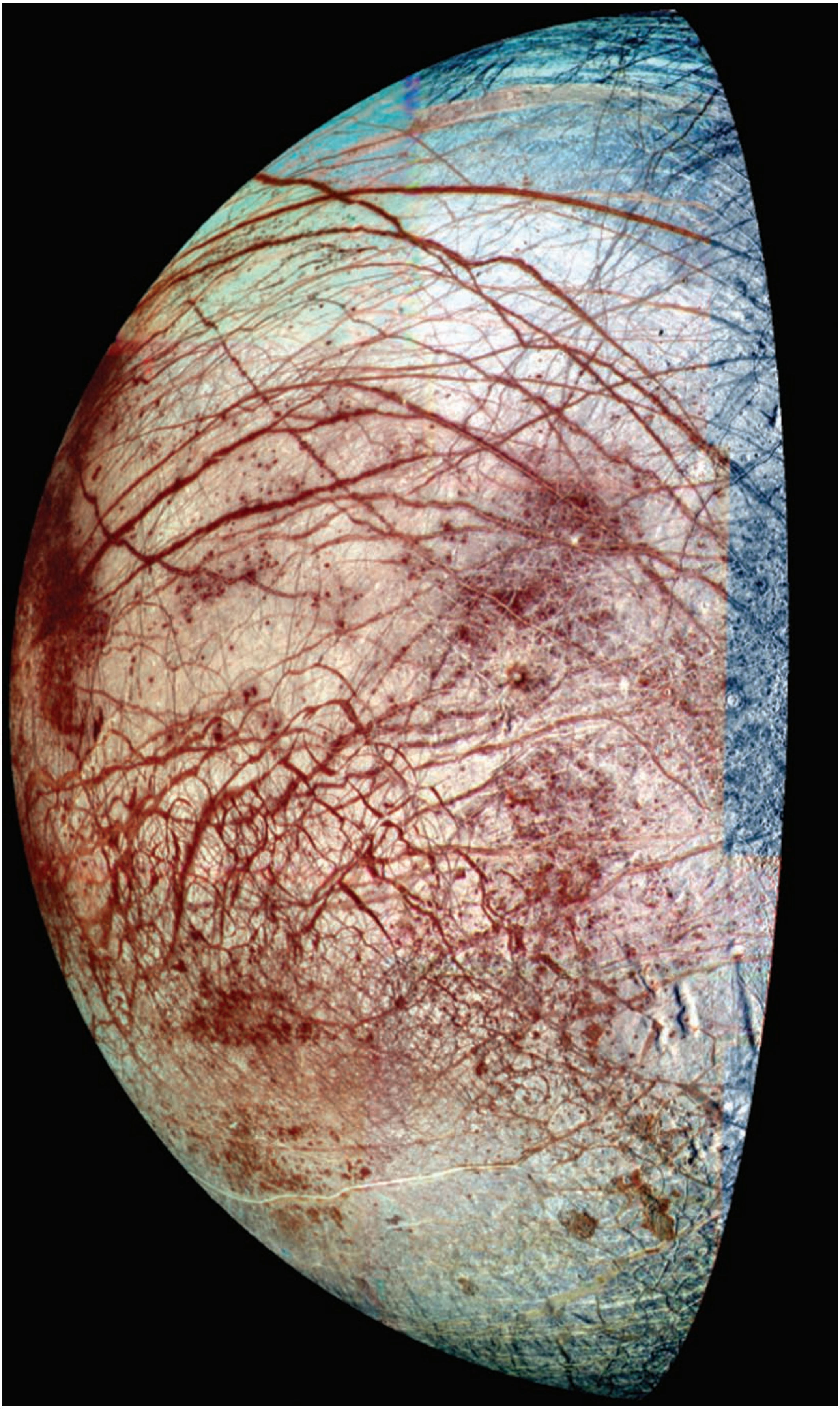
9.6 Jupiter's water moon Europa

Europa's bright, smooth icy complexion and young face

The smallest and yet brightest of the Galilean satellites, Europa, has a density comparable to that of rock, but its surface is as bright and white as ice. In fact, it is water ice! With surface temperatures of 110 kelvin or less, the water ice on Europa is frozen as hard and solid as granite.

Sunlight and charged particles cause some of the water ice to vaporize, and ultraviolet sunlight splits the molecules of water vapor into hydrogen and oxygen atoms. The hydrogen escapes into space, leaving behind a very tenuous atmosphere of oxygen.

Europa's surface is nearly devoid of impact craters, and there are no mountains or valleys on its bright smooth surface. No features extend as high as 100 meters, making Europa the smoothest moon or planet in the solar system.



The paucity of cratered impact scars indicates that Europa has a comparatively young surface, showing few signs of age. Since its surface must have been accumulating impact craters as time goes on, Europa must have been resurfaced in recent times, geologically speaking, probably in the past few hundred million years. Whatever is keeping Europa smooth is doing it from beneath the frozen crust, as eruptions of liquid water to the surface or flows of soft water ice. But it is unknown if the eruptions or flows are still occurring.

Long cracks, ice rafts and dark places on Europa

A veined, spidery network of long dark streaks marks Europa's young face, suggesting great inner turmoil. The fine lines run for thousands of kilometers, intersecting in spider-web patterns (Fig. 9.18). They give Europa a broken appearance that resembles a cracked mirror or an automobile window that has been shattered in some colossal accident. The dark lines are most likely deep fractures formed when that part of the ice cracked open, separated, and filled with darker, warm material seeping and oozing up from below. Dirty liquid water or warm dark ice has apparently welled up and frozen in the long cracks, producing the lacework of dark streaks (Fig. 9.19). They provide evidence for a young and thin, cracked and ruptured ice shell, probably moving slowly over the top of an ocean that is 100 kilometers or more deep.

The surface of Europa is fragmented everywhere, as if pieces of ice have broken apart, drifted away and then frozen again in slightly different places (Fig. 9.20). Large blocks of ice have floated like rafts across the moon's surface, shifting away from one another like moving pieces of a jigsaw puzzle. Some of them are tilted; others rotated out of place, like plastic toys bobbing in a bathtub. This shows that the ice-rich crust has been or still is lubricated from below by either slushy ice or maybe even liquid water.

The size and geometry of the ice floes on Europa suggest that internal heat has melted the ice just a few kilometers below the surface, producing an immense ocean that is hidden beneath the moon's frozen crust. The warmth and currents have broken the thin crustal ice into pieces that slide over the underlying watery slush. They resemble disrupted pack ice seen on Earth's polar seas during springtime thaws. But the thaw on Europa is coming from heat below, not from sunlight above.

Explosive ice-spewing volcanoes and geysers may erupt from the buried seas, reshaping the chaotic surface of the frozen moon and leaving dark scars behind. Extended dark regions may, for example, have formed when the subsurface ocean melted through Europa's icy shell, exposing darker material underneath, or when upwelling blobs of dark, warm ice broke through the colder near-surface ice.

Europa's subsurface sea of melted ice

Tidal distortions could explain how water ice has melted in the frigid environment near Europa. The satellite has a slightly eccentric orbit due to gravitational interactions with Io and Ganymede, which respectively revolve closer and farther away from Jupiter than Europa. Over the course of one trip around Europa's elongated path, Jupiter's strong gravity stretches and compresses the satellite in a process called tidal flexing. Frictional heat associated with similar tidal flexing melted the rocks inside Io, and it operates on Europa as well – to a smaller extent since Europa is further from Jupiter. But the warmth generated by tidal heating may have been, and may still be, enough to soften or liquefy some portion of Europa's icy covering, perhaps sustaining a subsurface ocean of liquid water.

The tidal flexing that warms Europa's interior may also crack the blanket of ice that traps the liquid water below. The varying distance of Europa from Jupiter causes the tides in the underground sea to rise and fall as much as 30 meters. The pressure of this continual, rhythmic in-and-out motion probably cracks the brittle crust apart.

Fig. 9.18 Broken ice on Jupiter's moon Europa This enhanced-color image shows subtle differences in the materials that cover the icy surface of Europa. Reddish linear crack-like features (*top*) extend for a thousand kilometers. They are caused by the tides raised in Europa by the gravitational pull of Jupiter. As the moon travels along its eccentric orbit the tides vary and fracture the thin, icy crust. The fractures probably open and become filled with a dirty slush from a possible ocean below. Mottled, reddish "chaotic terrain" exists where the surface has been disrupted and ice blocks have moved around. The red material at the fractures and chaotic terrain is a non-ice contaminant and could be salts brought up from a possible ocean beneath Europa's frozen surface. Also visible are a few circular features, which are small impact craters. Europa's surface has very few craters, indicating that recent or current geologic activity has removed the traces of older impacts. The paucity of craters, coupled with other evidence such as the red material, has led astronomers to propose that there might be an ocean of liquid water beneath Europa's surface. This view combines images from the *Galileo* spacecraft taken in violet, green and near-infrared wavelengths in 1995 and 1998. (Courtesy of NASA/JPL/U. Arizona.)

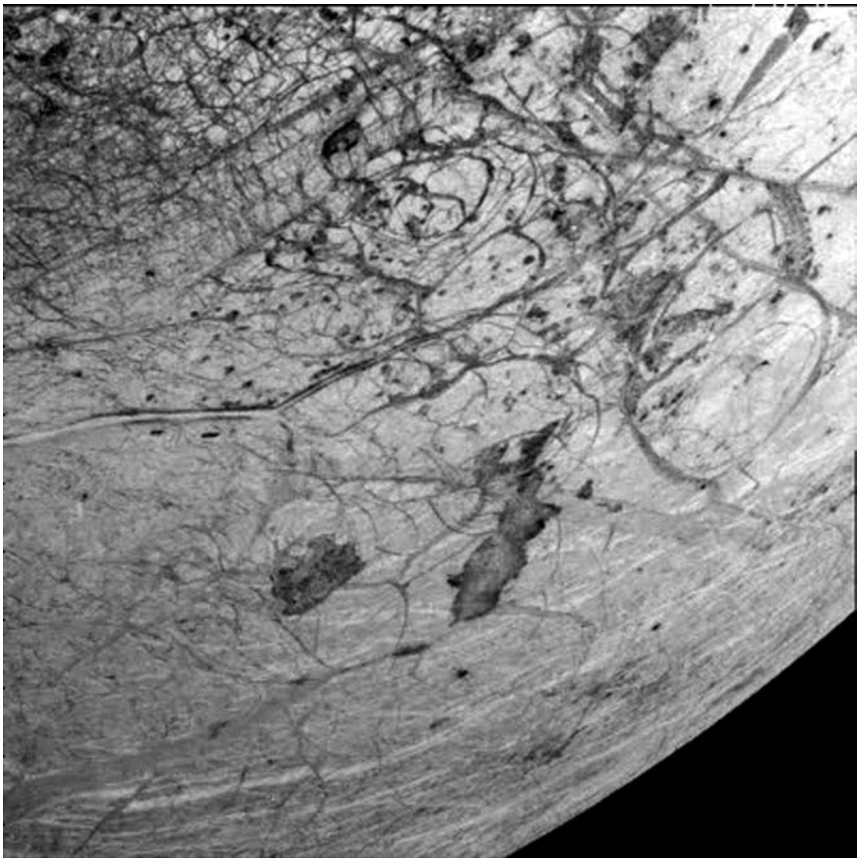


Fig. 9.19 Europa's frozen, disrupted surface Old impact craters are not visible on Jupiter's moon Europa. They must have been erased, perhaps by fresh ice produced along cracks in the thin crust or by cold glacier-like flows. The number of impact craters found on the bright, smooth surface indicates an age of approximately 100 million years. The thin, water-ice crust has undergone extensive disruption from below (*upper left*). Two irregular, chaotic dark features (*just below center*) were most likely formed when liquid water or warm ice welled up from underneath Europa's icy shell. These dark spots, technically called macula, are named Thera and Thrace after two places in Greece where Cadmus stopped in his search for Europa. This image, approximately 675 kilometers across, was taken from the *Galileo* spacecraft on 20 February 1997. (Courtesy of NASA/JPL.)

As Europa moves along its elliptical orbit and approaches Jupiter, the planet pulls more strongly on the moon, stretching its surface and cracking it open. Then as Europa recedes from Jupiter the gravitational stresses stop. By the time they begin again, Europa has a slightly different orientation, and the cracks start in a different direction, resulting in long, looping cracks that snake across the moon's scalloped surface.

Magnetic measurements from the *Galileo* spacecraft provide more evidence for an other-worldly ocean inside Europa. The satellite's magnetism changes direction as Jupiter's magnetic field sweeps by in different orientations to the satellite, owing to the tilt between the planet's rotation axis and magnetic axis. This means that the magnetic field at Europa is not generated in a core, but is instead induced by the passage of Jupiter's field through an electrically conducting liquid, such as salt water, beneath the ice. Although this evidence for a subsurface liquid ocean is indirect, it is the only indication that buried water is there now, rather than in the geological past.

So it is highly likely that Europa had liquid water near its surface at one time, and it might still be there. Gravity data tell us that the water moved to the top long ago, within an outer layer about 100 kilometers thick, and the cracked surface, floating icebergs, and changing magnetic field provide strong circumstantial evidence for internal seas just below the crust of ice. If the liquid water is still there, we can stretch our imagination and speculate that life might reside within its lightless depths (Focus 9.3).

Focus 9.3 Life in Europa's ice-covered ocean

The possibility of liquid water just below Europa's surface has led to speculation that life could have gained a foothold there. Tidal flexing might make the internal seas warm enough, and they would be wet enough; there might even be organic molecules down there. A global sea of liquid water could seethe with alien microbes hidden beneath Europa's gleaming ice-covered surface. After all, we know that the heat, minerals and chemical energy of underwater volcanoes on Earth's sea-floors sustain life in the dark without sunlight. Foot-long tubeworms and giant white clams thrive near the hot vents beneath the Earth's oceans, and microorganisms even live inside these volcanoes. This does not mean that there *is* life inside Europa, and there is no direct evidence for it, but it is an interesting speculation.

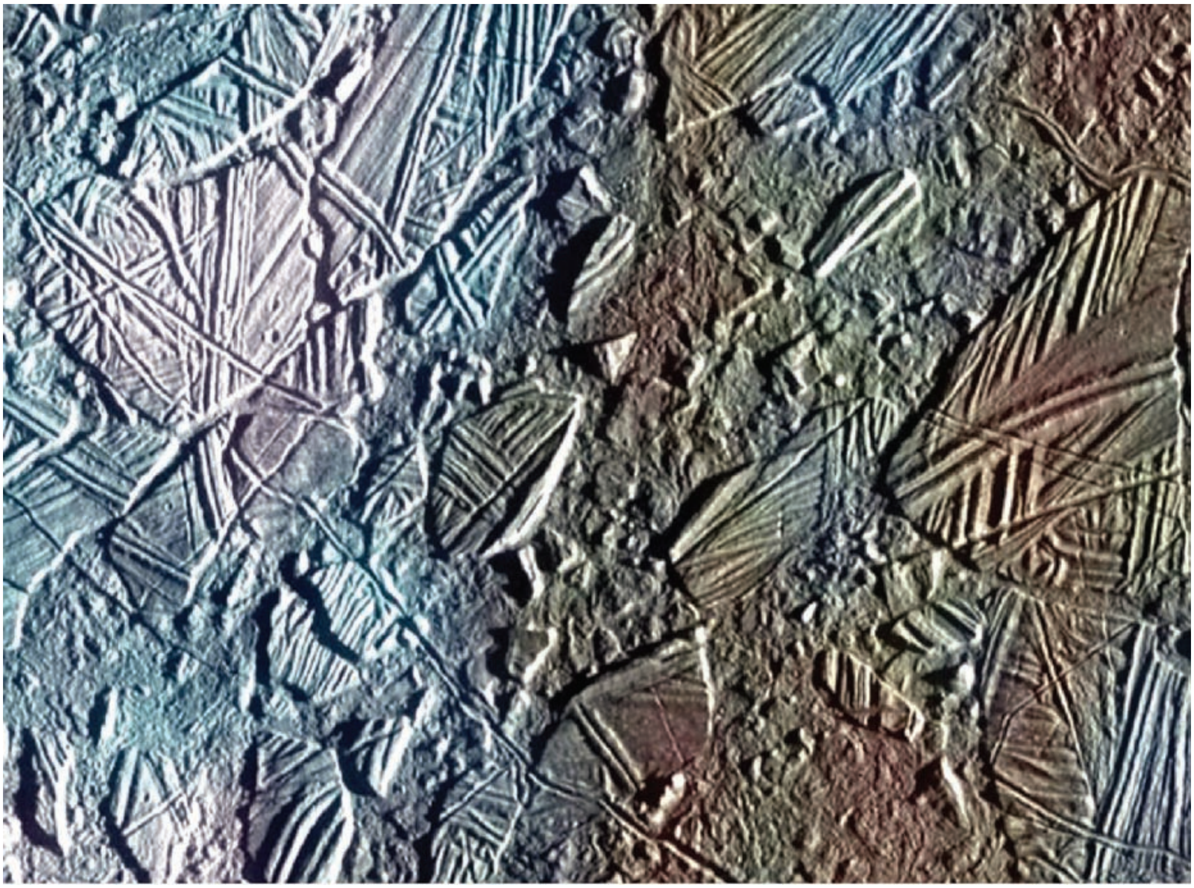


Fig. 9.20 Jupiter's moon Europa under stress This mosaic of the Conamara Chaos region on Europa indicates relatively recent resurfacing of the moon. Many sets of parallel and crosscutting ridges and fractures are detected. They are the frozen remnants of surface tension and compression, probably produced by heating and upwelling from below. The break-up and movement of the existing crust have formed irregularly shaped blocks of water ice. The blocks have been shifted, rotated, and even tipped and partially submerged within some lubricating material at the time of disruption, most likely soft ice or liquid water, or an ice-and-water slush below the surface. Some of the blocks have acted like rafts, separating and moving into new positions, somewhat like pack ice in the Earth's polar seas. The presence of young fractures cutting though this region indicates that the surface froze again after the resurfacing, forming solid, brittle ice. This image was taken from the *Galileo* spacecraft in February 1997. (Courtesy of NASA/JPL.)

9.7 Jupiter's battered moons, Ganymede and Callisto

Cratered, wrinkled Ganymede

Ganymede, the largest moon in the solar system, has a radius that exceeds that of the planet Mercury, although Ganymede is less dense and less massive than the planet. The satellite's mean mass density is so low that it must contain substantial quantities of liquid water or water ice. Like Europa, it is covered with ice. Ganymede's icy surface has experienced a violent history involving crustal fractures, mountain building and volcanoes of ice (Fig. 9.21).

Bright regions on Ganymede's surface contain sets of parallel ridges and valleys, termed grooved terrain, which looks like the swath of a giant's rake. The grooved terrain was most likely formed when the moon's water-ice crust expanded and stretched, cracking and rifting open as it was pulled apart. The crustal expansion might have happened when the satellite's rocks melted and moved into its interior while its water migrated to the top where it froze.

Sets of intersecting mountain ridges overlap and twist into each other. Some of the ridges cut across craters, while craters appear on other ridges (Fig. 9.22). Ganymede evidently experienced several epochs of mountain building. These crustal deformations may have continued for a billion years.

Water-ice volcanism played a role in creating the bright terrain on Ganymede. Prominent depressions were apparently flooded with liquid water or icy slush, and then froze

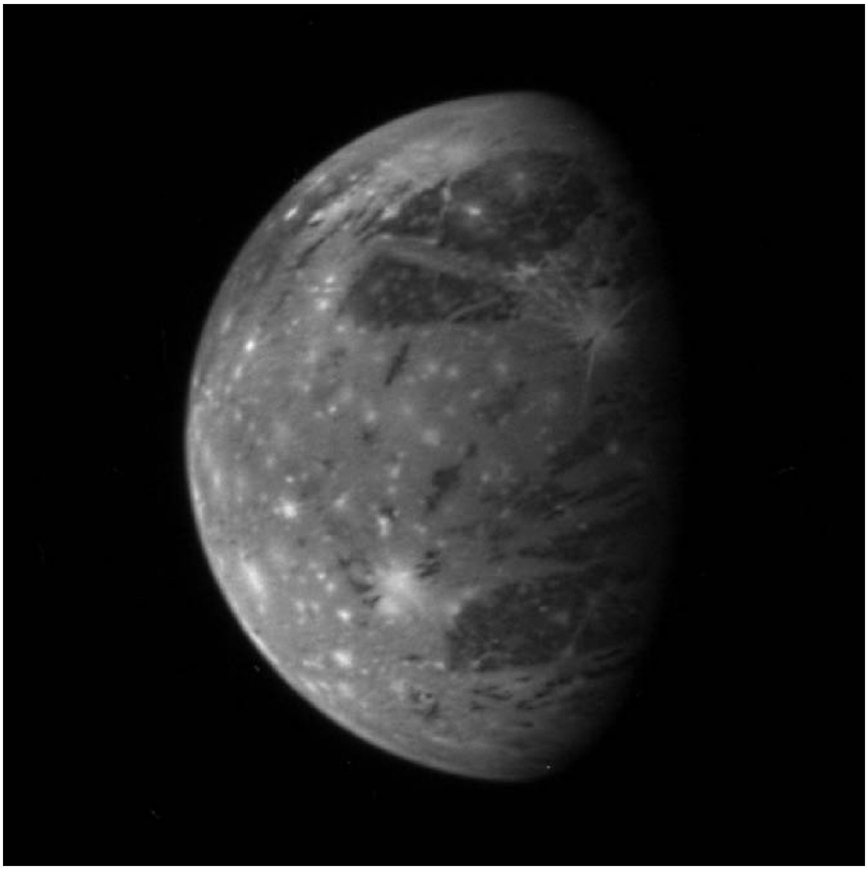


Fig. 9.21 Jupiter's large ice moon Ganymede Dark patches of ancient terrain are apparently frozen in swaths of bright, younger and translucent ice on Jupiter's satellite Ganymede, the largest moon in the solar system. More recent impact craters have splashed fresh, bright ice across the surface of the ice-covered moon. This image was taken on 27 February 2007 from the *New Horizons* spacecraft on its way to an expected encounter with Pluto in July 2015. (Courtesy of NASA/JHUAPL/SRI.)

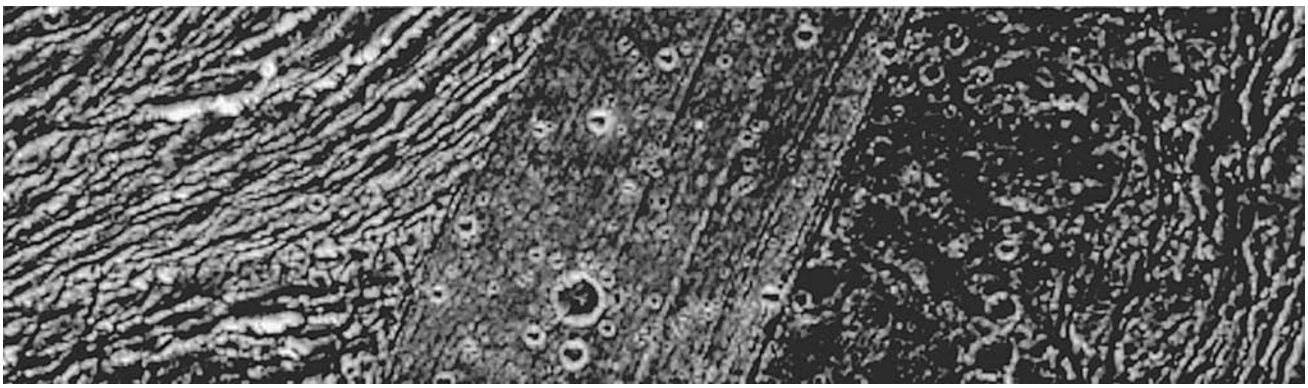


Fig. 9.22 Ganymede close up The bright icy crust on Jupiter's moon Ganymede contains both young and old terrain with bright grooves, caused by internal stress, and craters due to external impact. The youngest terrain (*center*) is finely striated and relatively lightly cratered. The oldest terrain (*right*) is rolling and relatively heavily cratered. The highly deformed grooved terrain (*left*) is of intermediate age. This image, approximately 89 kilometers across, was taken from the *Galileo* spacecraft on 20 May 2000. (Courtesy of NASA/JPL.)

into bright smooth bands that now cover much of the moon. Craters found in these areas indicate that this also happened early in the satellite's history, at least a billion years ago.

Darker regions on Ganymede are older and more heavily cratered. Ancient, densely packed impact craters on parts of Ganymede testify to the great age of the terrain, dating back 3 or 4 billion years. Some of the large polygonal blocks rise about a kilometer above the bright, grooved terrain, and look as if they have moved sideways for tens of kilometers along the moon's surface.

Ganymede: a moon with its own magnetic field

One of the major surprises of the *Galileo* mission was the discovery that Ganymede has its own intrinsic magnetic field. The moon is generating a magnetic dipole similar to those of many of the major planets, and roughly a thousandth of the strength of Earth's magnetic field. No other satellite now has such a magnetic field, but our Moon might have had one in the distant past.

Currents stirred inside Ganymede's large, dense molten iron core may have produced its self-generated magnetic field, but a molten core would cool down and the inner flows might last for only a million years or so. Scientists have therefore speculated that the satellite's orbit has shifted over time, and that strong tidal flexing once heated its interior more than it does now. If the moon moved in a closer or more eccentric orbit in the past, Jupiter's gravity would have squeezed it in and out by greater amounts, heating it up inside and generating a strong magnetic field that lingers today. The wrinkled, grooved terrain on Ganymede's icy surface might record this earlier period of intense heating. The hot core might still be cooling off, with internal currents that generate the magnetism seen today.

Since Ganymede's magnetic field is nestled within Jupiter's stronger and more extensive one, the giant planet's magnetic field sweeps past the moon. Magnetic readings taken from *Galileo* record the shifting magnetism that is thereby induced in Ganymede. Although this induced magnetic signature is much weaker than the satellite's intrinsic dipolar field, it suggests that a thick internal layer of salty water cause the induced field. Electric currents coursing through Ganymede's internal shell of salt-water can also contribute to its intrinsic magnetic field.

Callisto: an ancient, battered world

Remotest of the Galilean moons, Callisto is the third biggest moon in the solar system, after Ganymede and Saturn's moon Titan. Due to its larger distance from Jupiter, Callisto has had a much more sedate and peaceful history than the other Galilean satellites, with little sign of internal activity. It is a primitive world whose surface of ice and rock is the most heavily cratered in the solar system (Figs. 9.23, 9.24). Unlike nearby Ganymede, the moon Callisto has no grooved terrain or lanes of bright material, and it exhibits no signs of icy volcanism. So, Callisto is a long dead world unaltered by resurfacing since it formed and ancient impacts molded its face, a fossil remnant of the origin of the planets and their moons. In fact, with a surface age of about 4 billion years, Callisto has the oldest landscape in the solar system.

Yet, when seen close up by the *Galileo* spacecraft, there are indications of subdued, youthful activity on Callisto's surface. It is blanketed nearly everywhere by fine, mobile dark material, interrupted only where bright crater rims poke up through it (Fig. 9.25). Small impact craters are mostly absent, and those that are found sometimes appear worn down and eroded (Fig. 9.26). Thus, the smaller craters seem to have been filled in and degraded over time, perhaps by the dark blanket of debris that might have been thrown out by the larger impacts. Ice flows may have alternatively deformed and leveled many craters, because ice, which is rigid to sharp impact, can flow gradually over long periods of time, as glaciers do on Earth. The lack of small craters on Callisto might also be explained if the ancient population of impacting objects near the remote satellite had relatively few small objects when compared to the population near the Moon and Mercury.

Perhaps because Callisto is farther from Jupiter than the other Galilean moons, its ingredients are somewhat separated but still largely mixed, like a half-baked potato that is hard on the inside and soft on the outside. Unlike the other three Galilean satellites, Callisto has a homogeneous interior, without a dense metallic core, and it has no magnetic field of its own. But Callisto does have a crust of ice that may cover a subsurface ocean of liquid water.

Like Europa and Ganymede, the battered Callisto has a variable magnetic field, apparently generated by electrical currents as Jupiter's powerful field sweeps by. A shell of liquid water can explain the internal conductivity if it has the salinity of terrestrial seawater, but it would have to be deep enough inside the moon that the water could not rise to the surface, keeping it unaltered.

Since it does not participate in the orbital push and pull of Io, Europa and Ganymede, tidal flexing by Jupiter has not kneaded or heated Callisto inside. So Callisto's internal ocean can only be heated by radioactive elements. The lack of tidal flexing may also help explain the unwrinkled nature of Callisto's pockmarked face.

9.8 Jupiter's mere wisp of a ring

The rings of Saturn were discovered in the 17th century. About three centuries later, in 1977, several faint and unsuspected narrow rings were discovered about the planet Uranus. Jupiter was next to join the group of ringed planets, but this time the discovery was not a complete surprise. In 1974, the *Pioneer 11* spacecraft had encountered an unexpected reduction in the amount of high-energy charged particles when the spacecraft passed near Jupiter. Not much was made of this anomaly, although some

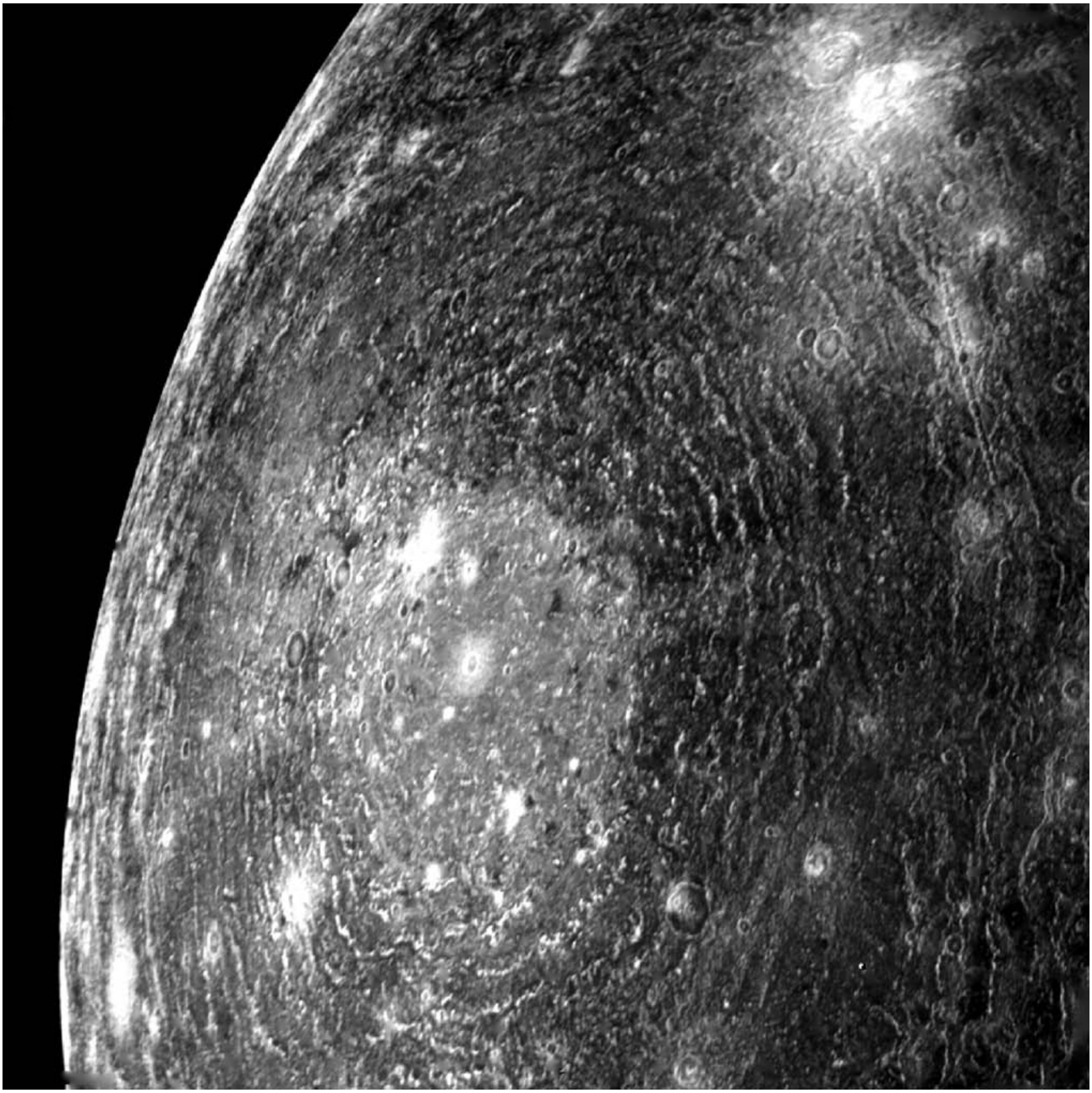


Fig. 9.23 Valhalla multi-ring basin on Jupiter's moon Callisto The icy crust of Callisto is as rigid as steel, and it therefore retains the scars of an ancient bombardment by impacting meteorites. An exceptionally powerful impact produced a multi-ring basin, named Valhalla after the home of the Norse gods. The extensive system consists of a light-floored central basin some 300 kilometers in diameter, surrounded by at least eight concentric mountainous ridges, which resemble ripples produced on a pond by a rock striking the water. The impacting object on Callisto apparently punctured the surface and disappeared. Today only the frozen, ghostlike ripples remain. The great number of rings observed around this basin on Callisto is consistent with the moon's low density and probably low internal strength. Although there are very few large craters, the rest of Callisto is pockmarked with smaller impact craters that are flat for their size, and many of them have bright rims that resemble clean water-ice splashed upon the dirtier surface ice. This image was taken from the *Voyager 1* spacecraft on 6 March 1979. (Courtesy of NASA/JPL.)

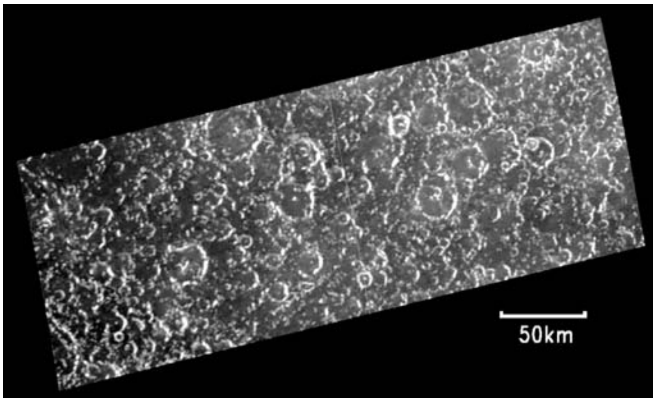


Fig. 9.24 Opposite the Valhalla impact on Callisto This close-up image was purposely taken on the opposite point, or antipode, of Callisto's huge Valhalla impact basin, created when an exceptionally large meteorite must have sent shocks into the moon's interior and waves rippling across the surface. The internal shocks ought to focus at the antipode, as they did for the Caloris impact on Mercury, creating a grooved and hilly terrain (also see Chapter 2, Fig. 2.23). The absence of such terrain at the Valhalla antipode suggests that Callisto has liquid water in its interior, which dispersed the seismic shocks. Magnetic field measurements also suggest that Callisto has a layer of liquid water deep below its surface. This image was taken from the *Galileo* spacecraft on 25 May 2001. (Courtesy of NASA/JPL/U. Arizona.)

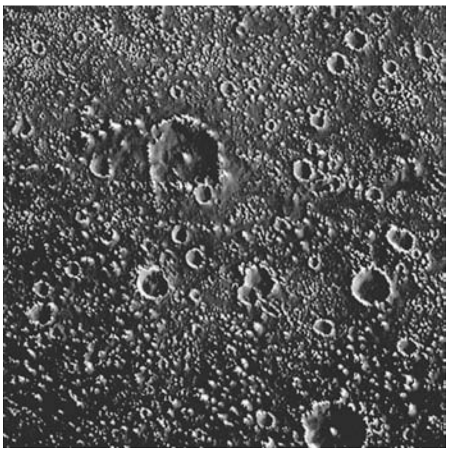


Fig. 9.25 Dark material and few small craters on Callisto A dark, mobile blanket of fine material covers Callisto's surface, sometimes collecting within crater walls. While Jupiter's moon Callisto is saturated with large impact craters, it has fewer very small craters when compared with the Earth's Moon and Mercury. One explanation is that the smaller craters have been filled by dark material that has moved down surface slopes. An alternative explanation for the paucity of little craters on Callisto is that there were fewer small impacting objects in its vicinity when compared with the amount within the inner solar system. This image, about 74 kilometers across, was taken from the *Galileo* spacecraft on 17 September 1997. (Courtesy of NASA/JPL.)

scientists thought that the falloff could be due to a previously unknown satellite or ring that blocked the energetic particles.

Finally, in 1979 – after much debate about the likelihood of finding a ring – a search was carried out with a camera on *Voyager 1*, and a narrow faint belt of material was found encircling the planet in its equatorial plane near the same distance that the energetic particles had disappeared. The ring was not previously observed from Earth because it was too faint and close to the bright planet. Since its discovery, Jupiter's main ring has been detected by Earth-based telescopes sensing infrared radiation, from the *Hubble Space Telescope*, and more fully explored by the inquisitive eyes of the *Galileo* spacecraft (Figs. 9.27, 9.28).

When the retreating *Voyager 1* camera looked back at the shadowed side of Jupiter, the main ring became brighter. The Sun was backlighting the ring's tiny particles, making them shine brightly. This behavior is typical of very small particles that scatter light in the forward direction, like tiny salt grains on the windshield of an automobile, the smoky haze in some movie theaters, or the condensation trails of airplanes.

The size of the ring particles can be inferred from the way they scatter light, and the conclusion is that they are a few millionths of a meter across or about the same size as flour dust or the grains of pollen. The particles that make up cigarette smoke and the hazes in the Earth's atmosphere have similar sizes. Numerous, larger ring particles would have reflected sunlight, making the rings appear brightest when approaching them, which did not happen.

Voyager 1 and *2* viewed a three-ring system around Jupiter, consisting of a flattened main ring, an inner extended cloud-like ring, called the halo, and a third outer ring, known as the gossamer ring because of its transparency. Observations from the *Galileo* spacecraft in 1996 and 1997 showed that the gossamer ring consists of two parts, one embedded in the other.

Jupiter's insubstantial ring system is practically made of nothing at all, no dustier than a typical living room. And the individual dust particles only reside temporarily in Jupiter's rings, just as the dust in the air of your room settles onto the room's furniture and bookshelves.

The dust particles in Jupiter's rings don't stay in the rings forever. They can last no more than a few thousand years before being tossed out of the ring plane or spiraling down into Jupiter's upper atmosphere and eventually disappearing. Given this short lifetime, the fine particles are probably replenished if the Jovian rings are permanent features.

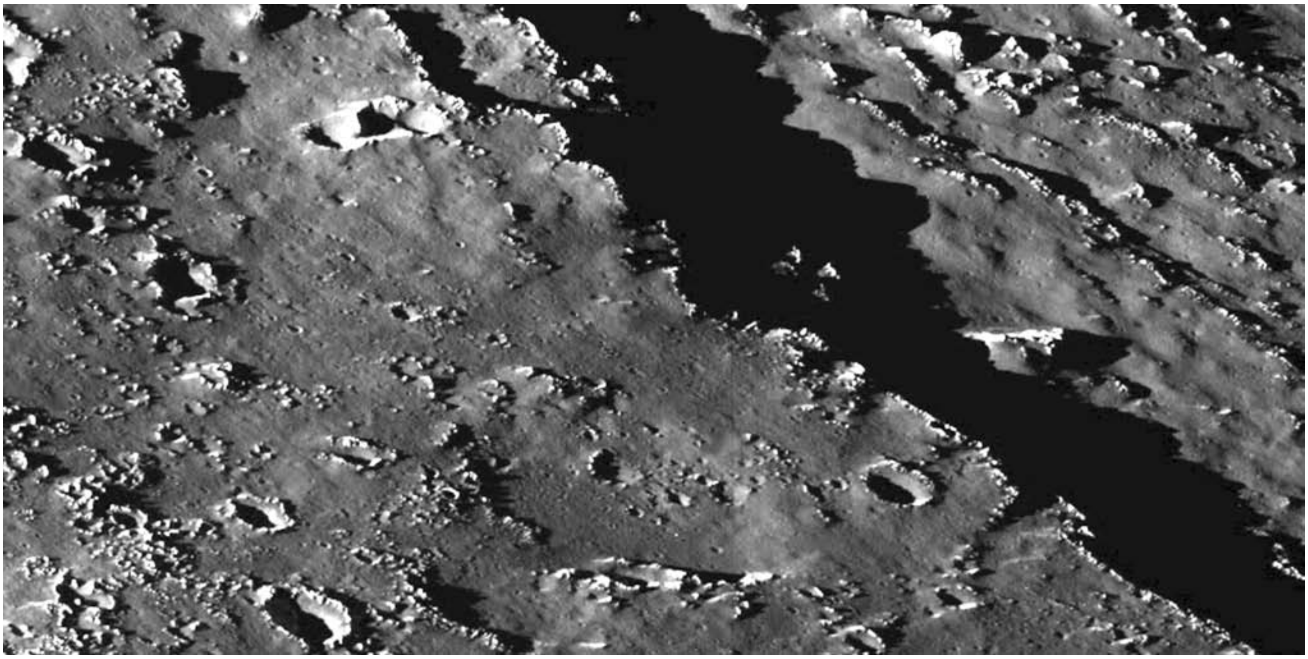


Fig. 9.26 Structures within the Valhalla impact basin on Jupiter's moon Callisto This mosaic of two images shows an area within the Valhalla region on Jupiter's outermost large moon, Callisto. Numerous impact craters, ranging in size from 155 meters to 2.5 kilometers across, are seen in the mosaic, which is approximately 33 kilometers wide. There is an unexpected absence of small craters and the ragged rims of the existing craters suggest an erosion process, which is believed to be due to sublimation of volatiles from the surface ice. A prominent dark fault scarp, or cliff, and smaller parallel ridges are also pictured (*top right*). These images were obtained on 4 November 1996 from the *Galileo* spacecraft. (Courtesy of NASA/JPL/ASU.)

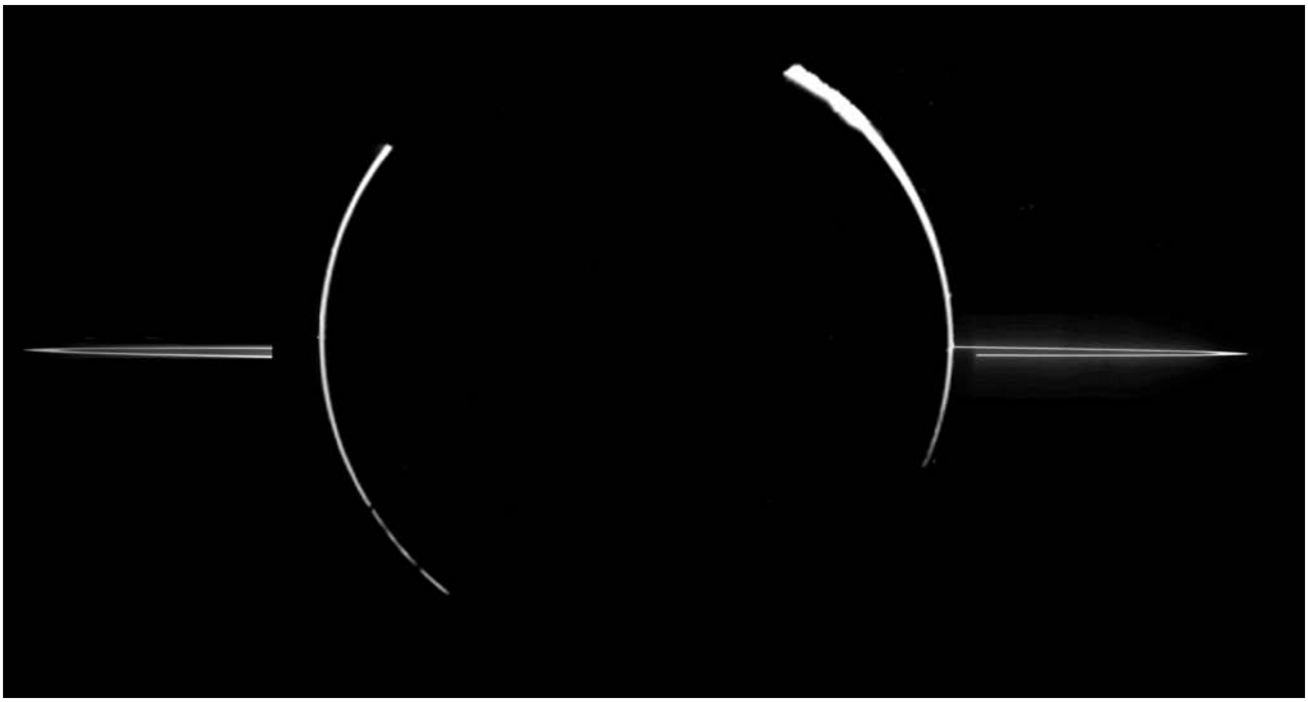


Fig. 9.27 Jupiter's rings The upper atmosphere of Jupiter and the planet's main ring can be seen when the Sun is behind the planet, and an imaging spacecraft is in Jupiter's shadow peering back toward the Sun. In such a configuration, very small dust-sized particles are accentuated so both the ring particles and the smallest particles in the upper atmosphere of Jupiter are highlighted. It is somewhat like looking back at a movie projector in a dusty theater or at a bright light in a smoky room, which permits you to see the dust or smoke in the air. The small particles in Jupiter's rings are believed to have human-scale lifetimes, and must be continuously replenished if the ring persists. In addition to the flat, main ring, Jupiter's ring system includes a toroidal halo interior to the main ring and a gossamer ring, which only become visible when images are overexposed with respect to the main ring. This image was taken on 9 November 1996 from the *Galileo* spacecraft. (Courtesy of NASA/JPL/Cornell U.)

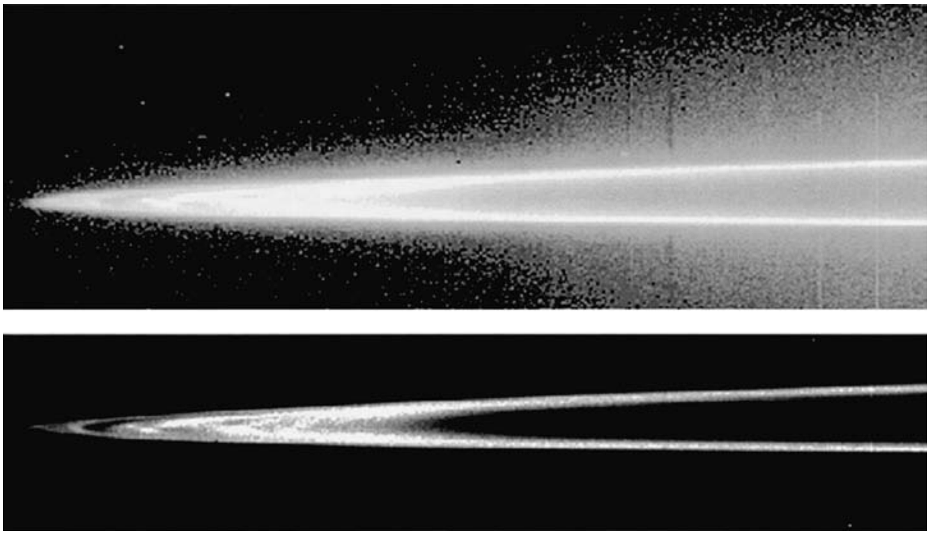


Fig. 9.28 Jupiter's main ring and halo Jupiter's bright, flat main ring (*bottom*) is a thin strand of material encircling the planet with an outer radius of 128 940 kilometers, or about 1.8 times Jupiter's radius, located very close to the orbit of the giant planet's small moon Adrastea, at 128 980 kilometers. The brightness of the main ring drops markedly very near the orbit of another moon, Metis. A faint mist of particles, known as the ring halo, surrounds the main ring and lies above and below it (*top*). The vertically extended halo is unusual for planetary rings, which are normally flattened into a thin plane by gravity and motion. The halo probably results from the "levitation" of small charged particles that are pushed out of the main ring plane by electromagnetic forces. These images were obtained from the *Galileo* spacecraft on 9 November 1996 when it was in Jupiter's shadow, looking back toward the Sun. The rings of Jupiter proved to be unexpectedly bright when seen with the Sun behind them, just as motes of dust or cigarette smoke brighten when they float in front of a light. A third gossamer ring, which consists of two components, is not shown here; it lies beyond the main ring, at greater distances from Jupiter. (Courtesy of NASA/JPL.)

They might be supplied by dust that is blasted off Jupiter's four small, innermost moons by interplanetary meteoroids, the fragments of comets and asteroids. The meteoroids are drawn in by Jupiter's very strong gravity, which also greatly increases their speed. When the high-velocity cosmic meteorites slam into one of the small inner moons, they create a dust cloud. The dust is thrown off at such high velocity that it escapes the moon's relatively small gravitational pull, orbiting Jupiter and contributing to one of its rings.

The outer edge of the main ring lies just inside the orbit of the tiny moon Adrastea, just 15 kilometers in size and too small to be seen from Earth. It was discovered by the *Voyager* spacecraft, as was another tiny moon, named Metis, which is embedded near the bright midpoint of the main ring. The dust generated by meteorite impact on Adastea and Metis can easily escape the small gravity of these moons, accounting for the dense accumulation of particles in the main ring. Some of the microscopic particles are small enough that if they are slightly charged then electromagnetic forces can overpower the effects of Jupiter's gravity, pumping them into the inner halo that is seen above and below the main ring.

The two much fainter, and wider, gossamer rings, which are more distant from Jupiter than the main ring, lie just inside the orbits of the small moons Amalthea and Thebe. Detailed observations from the *Galileo* spacecraft indicate that dust particles knocked off these two satellites feed the two gossamer rings, with a thickness that corresponds to each satellite's elevation above the planet's equatorial plane.

Jupiter has more than 63 moons, and most of them have been discovered using ground-based telescopes instead of by close-up spacecraft observations. Large charge-coupled device cameras and computers have been attached to the world's largest telescopes and programmed to detect anything that moves near Jupiter. The new moons are all small, between 1 and 10 kilometers across, and most of them move along retrograde orbits, in the opposite direction to the major satellites of Jupiter. The numerous small moons were probably captured by Jupiter, perhaps when its former, more extended atmosphere slowed the passing moons down enough to be held in the giant's gravitational embrace.

Altogether, Jupiter's dark rings are as wide as Saturn's yet invisible from terrestrial telescopes. Unlike Saturn's rings, which are made of bright, highly reflective

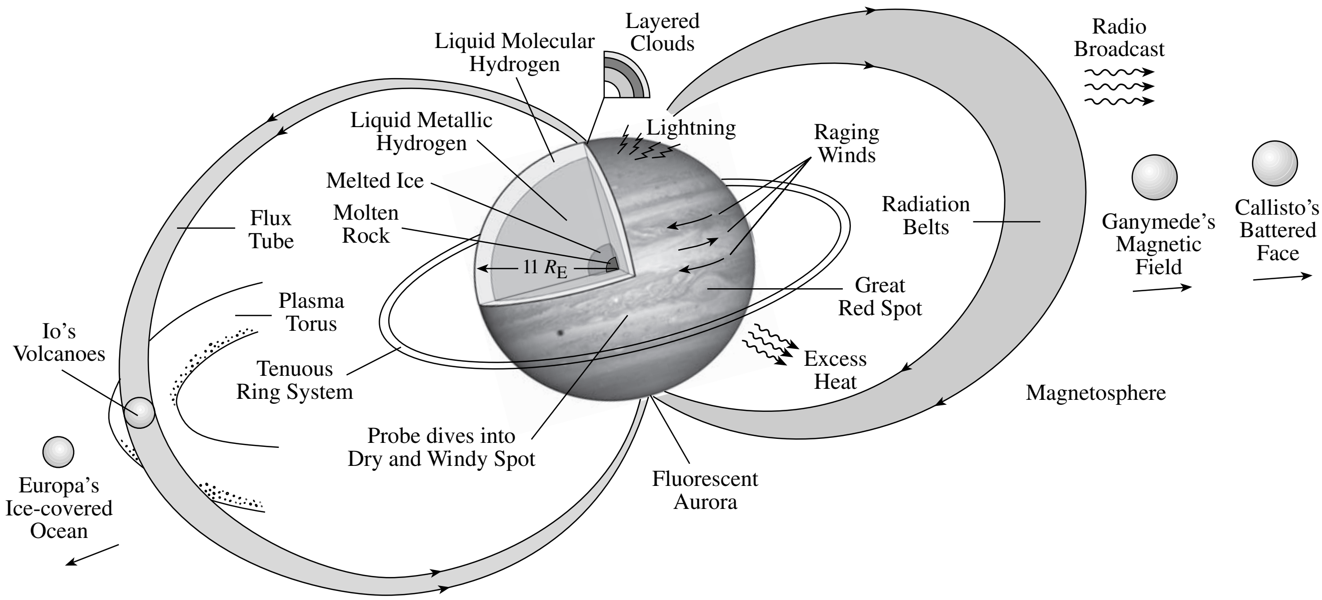


Fig. 9.29 Summary diagram

water ice, in pieces as large as houses, Jupiter's rings consist of fine dust grains, so dark that they reflect barely 5 percent of the sunlight that hits them. They are also spread so thin that Jupiter's rings are almost transparent. So this naturally brings us to Saturn, lord of the rings.

10 Saturn: lord of the rings

- Saturn has the lowest mass density of any planet in the solar system, low enough for the planet to float on water, and this means that Saturn is primarily composed of the lightest element, hydrogen.
- Saturn's rapid rotation has pushed its lightweight material out in the planet's middle, creating the most pronounced equatorial bulge of any planet.
- Saturn is just a great big liquid drop, covered by a thin atmosphere of gas, slightly smaller than Jupiter and less than a third of its mass.
- Liquid hydrogen is compressed inside Saturn's depths to form an electrically conducting, liquefied metal.
- There is no solid surface anywhere inside Saturn, though it might have a core of melted ice and molten rock that is about ten times as massive as the Earth.
- Saturn radiates almost twice as much energy as it receives from the Sun, and most of the planet's excess heat is generated by helium raining down into its liquid metallic hydrogen core.
- Saturn's rings are completely detached from the planet and separated from each other.
- The rings of Saturn are not solid, but instead composed of innumerable small water-ice particles and larger chunks of water ice.
- The icy constituents of Saturn's main A, B and C rings are as big as hailstones, snowballs and even icebergs; there are more smaller ones, but the big kind supply most of the ring mass.
- The total mass of the rings of Saturn is comparable to that of its medium-sized satellite Mimas, which is 396 kilometers across.
- Saturn has a retinue of diffuse, tenuous, and nearly transparent rings, designated the D, E, F and G rings, that are most likely composed of microscopic ice crystals, smaller than snowflakes and about the size of the dust in your room.
- Two small moons confine the edges of Saturn's narrow F ring and shepherd the ring particles between them.

- The icy material in the prominent rings of Saturn has been marshaled into thousands of individual ringlets, resembling ripples on a pond, but with circular, oval and even spiral shapes.
- Gravitational interaction with nearby external satellites can sculpt Saturn's ring material into numerous ringlets and produce waves in it.
- Small moons embedded within Saturn's rings can sweep out gaps, keeping them open and also sharpening their edges.
- The gaps within Saturn's rings are not completely empty; the Cassini Division contains about 100 ringlets.
- Enigmatic dark spokes stretch in the radial direction across Saturn's rings, moving at constant speed regardless of distance from the planet, in apparent violation of the laws of gravity.
- Saturn's dark ring spokes consist of microscopic dust-sized particles that may become electrically charged and levitated above the larger ring particles. They might then be swept around Saturn by its rotating magnetic fields.
- Planetary rings lie closer to a planet than its large satellites, within the Roche limit where the planet's tidal forces will rip a large satellite to pieces and prevent small bodies from coalescing to form a larger moon.
- The rings of Saturn could have formed when a moon was pulled toward the planet by tidal forces and eventually ripped apart.
- Small moons embedded in the rings of Saturn might sustain them.
- Saturn's relatively small moon Enceladus emits jets of ice particles, powdery snow, water vapor and organic compounds, which vent from warm fissures, known as tiger stripes, in the moon's south polar crust.
- The active jets on Enceladus suggest that tidal effects may make the moon hot inside. The tiger-stripe fractures could rub against each other, creating heat, or open to expose explosive ice to the vacuum of space. The interior ice might be melted into subsurface seas of liquid water containing organic chemical elements.
- Saturn's largest moon, Titan, is a planet-sized world with a substantial atmosphere whose surface pressure is about 1.5 times the air pressure at sea level on Earth.
- Titan's atmosphere is composed of 98.4 percent molecular nitrogen, nearly 1.6 percent methane, and trace amounts of other hydrocarbons; so nitrogen molecules are the main constituents of Titan's atmosphere, as they are in the Earth's air.
- Hazy smog, composed of complex organic molecules, envelops Titan's atmosphere and hides its surface from view.
- The *Huygens Probe* touched down on the surface of Titan on 14 January 2005, detecting methane rainfall and dark narrow riverbeds on the way down. The probe landed at equatorial latitudes, on a damp, moist riverbed littered with pebbles that were apparently rounded by flowing liquid.
- Radar pulses from the *Cassini* spacecraft in orbit about Saturn have seen through the haze that shrouds Titan, revealing long, deep, meandering channels on Titan's surface, which resemble terrestrial rivers but are attributed to flowing methane or ethane rather than water.

- The *Cassini* radar instrument has imaged dark, flat, smooth places with shore-like boundaries. They have been attributed to large lakes of liquid methane and ethane, and the spectral signatures of liquid ethane have been detected in at least one of them.
- Seasonal variations might account for the fact that about 20 more lakes were found at high northern latitudes on Titan than high southern ones. Clouds may rain methane during winter in the north, when southern lakes are evaporating in the local summer.
- Vast dunes accumulate near Titan's equator, shaped by strong winds blowing east to west. Unlike Earth's sand dunes, Titan's dunes are thought to be composed of organic material that has rained down from its smoggy skies.
- The dunes and lakes on Titan may contain hundreds of times more hydrocarbons than all the oil and gas reserves on Earth.
- Saturn has six mid-sized icy moons that retain impact craters dating back to their early history; some of them exhibit signs of internal activity and ice volcanism. Impacting objects almost broke the moons Mimas and Tethys apart.
- A number of unique small, irregularly shaped moons revolve around Saturn with remarkable orbits. The co-orbital moons have almost identical orbits, the Lagrangian moons share their orbit with a larger satellite, and the shepherd moons confine the edges of rings.
- Saturn's mid-sized moon Hyperion is so light that it must be about half-filled with empty spaces, and it tumbles chaotically along its orbit with no definite rotation period or orientation in space.
- The enigmatic moon Phoebe moves around Saturn in the opposite, retrograde direction to the planet's other mid-sized satellites. Phoebe has sharp-edged craters and a varying brightness that suggest thin, dark, surface deposits overlying bright water ice.
- An enormous, exceptionally distant, wide and diffuse ring of Saturn is apparently replenished by micrometeorite impacts with Phoebe; the ejected material might also move in toward Saturn, striking the next innermost moon Iapetus and accounting for the dark side of its two-faced surface.

10.1 Fundamentals

Majestic Saturn, the sixth planet from the Sun, was the most distant world known to the ancients, and it moved least rapidly around the zodiac. The Greeks identified the planet with Kronus, the father of Zeus, while the Romans named the planet Saturn after their god of sowing. Both the Greeks and the Romans associated Saturn with the ancient god of time, who later became Father Time.

You can see Saturn's oblong, golden disk with a small telescope, girdled by its beautiful rings, unattached to the globe. They set Saturn apart from all the other planets. Even though we now know that all four of the giant planets possess ring systems of some kind, Saturn's rings easily outclass the others.

Saturn's realm has been scrutinized in close-up detail from the *Voyager 1* and *2* flyby spacecraft, in 1980 and 1981 respectively, and from the orbiting *Cassini* spacecraft from 2004. The two *Voyagers* sent back marvelous images of Saturn's yellow-brown clouds, magnificent rings and large icy satellites (Fig. 10.1). The instruments aboard *Cassini* have provided captivating images as beautiful as a work of art (Fig. 10.2). The *Hubble Space Telescope* has also been returning high-resolution images of Saturn's rings and banded atmosphere, and ground-based telescopes have been monitoring their infrared heat radiation.

Saturn's orbital radius is 9.5 times the radius of the Earth's orbit, and it takes slightly more than 29 Earth years for Saturn to complete one revolution around the Sun. Perhaps because of its remote orbit and slow motion, the planet's name has been adopted for the word "saturnine", to describe a cool and distant temperament.

At its large distance from the Sun, the ringed planet and its satellites receive only about 1 percent as much sunlight

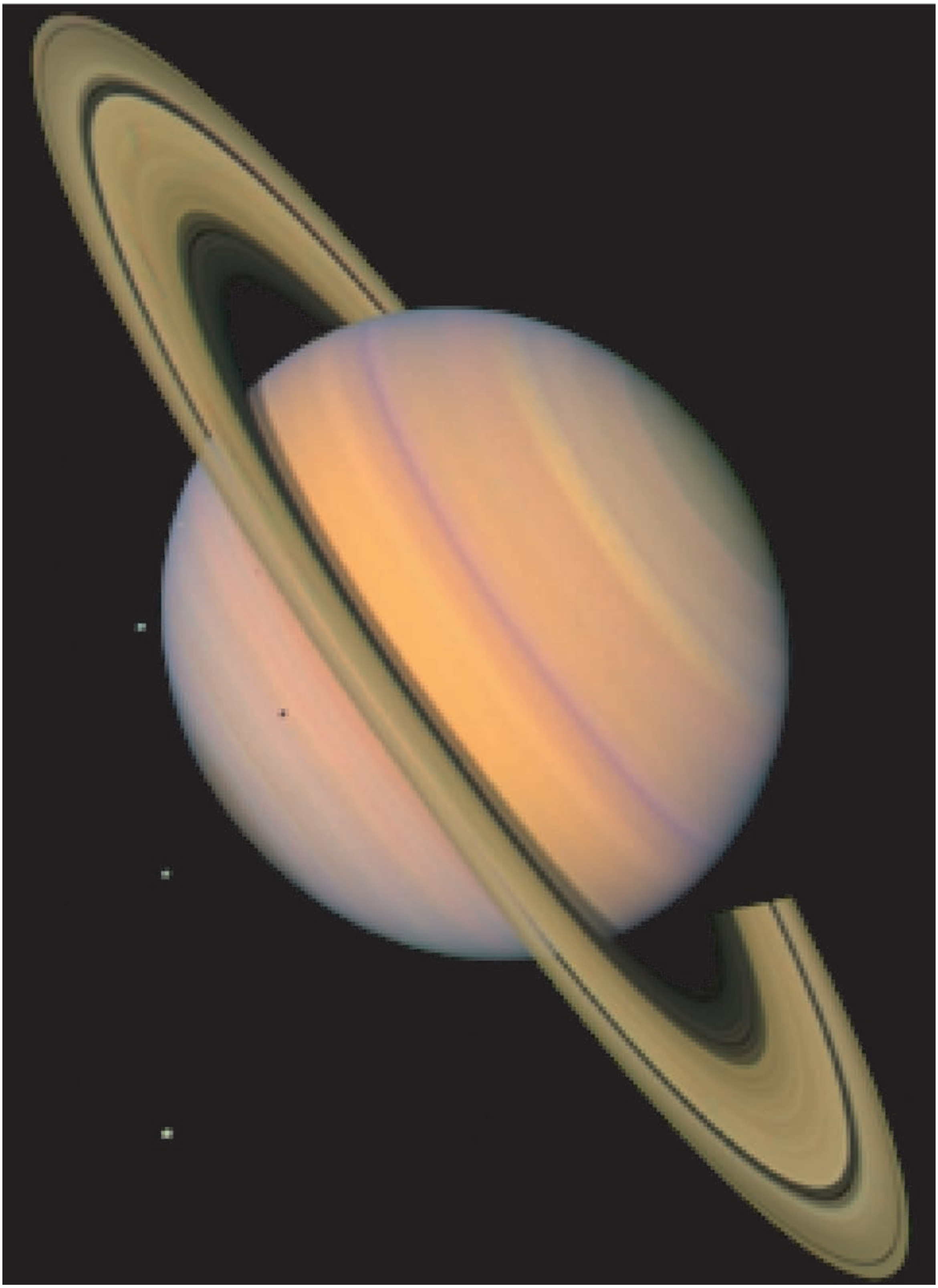


Fig. 10.1 Saturn's realm The magnificent rings of Saturn encircle the planet, never touching its cloud tops. From the outside in, there are the bright A and B rings separated by the Cassini Division. The narrow Encke Gap in the outer A ring is also visible, as is the dark C ring nearest to the planet. The yellow-brown atmosphere of Saturn, shown here in enhanced color, has a banded structure, but it lacks Jupiter's bright zones and belts. Three icy satellites (Tethys, Dione and Rhea) are visible as small white spots against the darkness of space (*bottom left*), and another smaller satellite (Mimas) is visible just above them on Saturn's bright edge or limb. Small black round shadows cast by Mimas and Tethys are visible on Saturn's disk, and the planet blocks light from getting to the rings at the right just outside the planet. Because of its rapid spin, Saturn has an oblong, egg-like shape, flattened at the poles and extended at the equator. This *Voyager 2* image was acquired on 4 August 1981. (Courtesy of NASA/JPL.)

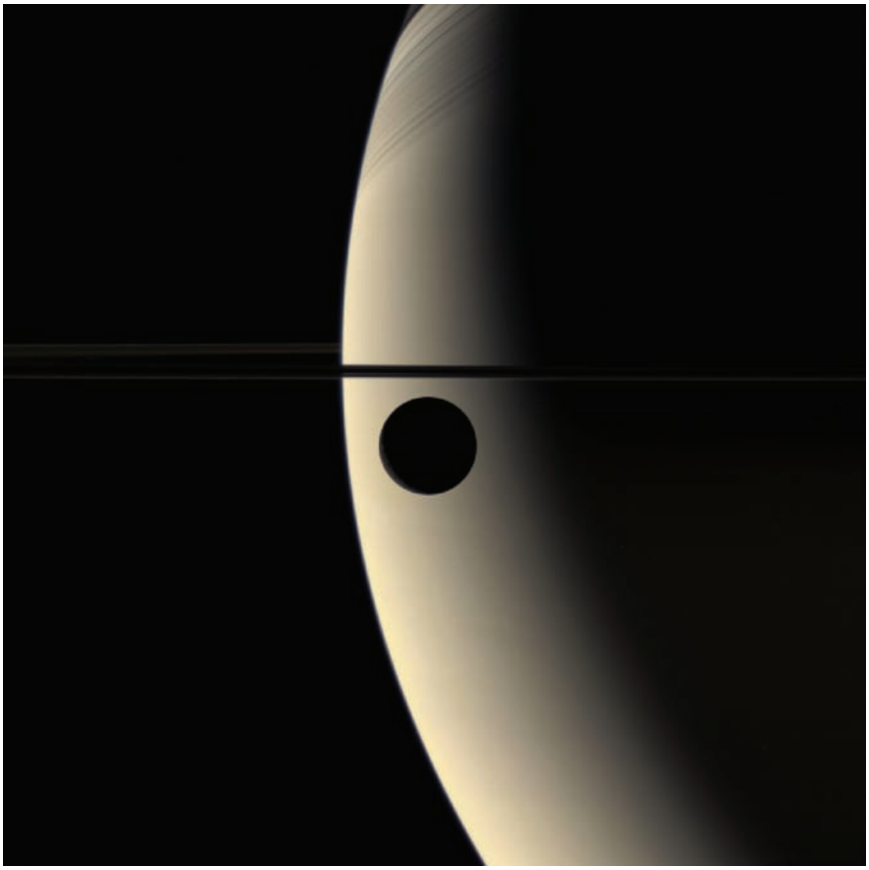


Fig. 10.2 Rhea passes in front of Saturn The moon Rhea glides above the featureless, golden face of Saturn. This view looks down onto the unlit side of Saturn's normally impressive rings, which are visible here only as a thin line. This image was acquired from the *Cassini* spacecraft on 21 March 2006. (Courtesy of NASA/JPL/SSI.)

and solar heat as the Earth does. The surfaces of many of Saturn's satellites are therefore covered with water ice. And even though Saturn generates some of its own heat, its cloud tops have a temperature of only 95.0 kelvin.

Saturn is the second largest planet in the solar system, overshadowed only by Jupiter. The radius of Saturn, without the rings, is about four-fifths the radius of Jupiter and slightly more than 9 times the radius of the Earth.

The volume of Saturn is great enough to encompass 764 Earth-sized planets. But Saturn's mass is only 95 times greater than the Earth's mass, so the giant planet must be composed of material that is much lighter than rock and iron, the primary ingredients of the Earth.

From Saturn's mass and volume, we calculate its average mass density, or bulk density, to be only 687 kilograms per cubic meter, the lowest of any planet and less than that of liquid water. If Saturn were placed in a large enough ocean of water, it could float. It has a low average density because it is mainly composed of the lightest elements, hydrogen and helium, in the gaseous and liquid states.

Saturn rotates with a day of only 10.6562 hours – only 44 minutes longer than Jupiter's rotation period of 9.9249 hours. This is the rotation period of Saturn's magnetic field that is anchored inside the planet, and it is inferred from the observed periodic modulation in Saturn's radio emission, generated in the spinning magnetic fields. The visible clouds spin at different speeds, faster at the equator and slower at the poles. Like Jupiter, the rapid rotation has stretched Saturn's clouds around the planet in bands, but with a buttery color and weather patterns that are harder to see.

10.2 Winds and clouds on Saturn

Ferocious winds circle Saturn in jet streams that reach 500 meters per second, almost four times the speed of Jupiter's fastest winds and ten times hurricane force on the Earth. The dominant winds on Saturn blow eastward, in the same direction as the planetary rotation, at almost all latitudes, with the most powerful nearest to the equator. Reversals in wind direction are only found near Saturn's poles, where the clouds counter-flow in the eastward and westward direction. They form banded belts and zones similar to those observed almost everywhere on Jupiter. At each pole, the winds spiral downward in a high-speed vortex larger than the Earth, like the

Table 10.1 Physical properties of Saturn[a]

Mass	$5.683\,19 \times 10^{26}$ kilograms = 95.184 M_E
Equatorial radius at one bar	60 268 kilometers = 9.46 R_E
Mean radius at one bar	58 232 kilometers
Mean mass density	687.1 kilograms per cubic meter
Sidereal rotation period	10 hours 39 minutes 22.3 seconds = 10.6562 hours
Sidereal orbital period	29.4475 Earth years
Mean distance from Sun	1.4294×10^{12} meters = 9.539 AU
Age	4.6×10^9 years
Atmosphere	97 percent molecular hydrogen, 3 percent helium
Energy balance	1.79 ± 0.10
Effective temperature	95.0 kelvin
Temperature at one-bar level	134 kelvin
Central temperature	13 000 kelvin
Magnetic dipole moment	600 D_E
Equatorial magnetic field strength	0.22×10^{-9} tesla or 0.72 B_E

[a] The symbols M_E, R_E, D_P, B_E denote respectively the mass, radius, magnetic dipole moment and magnetic field strength of the Earth. One bar is equivalent to the atmospheric pressure at sea level on Earth. The energy balance is the ratio of total radiated energy to the total energy absorbed from sunlight, and the effective temperature is the temperature of a black body that would radiate the same amount of energy per unit area.

deep eye of a cosmic hurricane complete with multiple thunderstorms.

Despite its raging winds, Saturn lacks the dynamic and colorful storm clouds of Jupiter. Stormy weather on Saturn is apparently masked by an upper deck of dirty, smog-coated particles that give the planet a pastel, butterscotch hue at the visible wavelengths we see with our eyes. Jupiter, being warmer than Saturn, has less of this smoggy haze, and its cloud features are more distinct at visible wavelengths.

The clouds and hazes in Saturn's atmosphere are perceived in greater detail at ultraviolet and infrared wavelengths (Figs. 10.3, 10.4). Infrared images can also detect the heat radiated by the warmer parts of Saturn's atmosphere.

On rare occasions, a gigantic storm cloud of fresh, clean, white ammonia ice warms up, rises and punches through the opaque upper cloud deck, somewhat like the upwelling of warmer air in terrestrial thunderheads. The swirling, white equatorial ovals have been recorded by ground-based telescopic observers, but only a few times in the past two centuries. High-resolution images from the *Hubble Space Telescope* are clear enough to pick out the details of large storm systems and to record minor storms that look like bright cloud features. Observations from the *Cassini* spacecraft suggest that the rotating storms power the jet streams.

Instruments aboard *Cassini* have also detected intense ammonia-ice blizzards, apparently energized from below, where lightning has been observed and the clouds include water. The powerful storms produce lightning flashes that are as bright as the brightest ones seen on Earth, and the radio noise produced by Saturn's lightning bolts produce as much static at the spacecraft as terrestrial lightning causes in an AM radio on Earth.

When *Voyager 2* passed behind Saturn, its homebound radio signals penetrated the upper atmosphere, and alterations in these transmissions have been used to deduce the pressure and temperature below the clouds (Fig. 10.5). Because there is no solid surface directly below the clouds, altitudes are referred to the level in the atmosphere where the pressure is equal to 0.1 bars, or one-tenth the sea-level pressure on Earth. This is the approximate level where the temperature bottoms out, at about 82 kelvin, and the obscuring veil of haze may be formed.

Under the assumption that Saturn's gas mixture is in chemical equilibrium, with a uniform composition like that of the Sun, ammonia is expected to condense and form clouds at about 100 kilometers below the reference level, where the pressure has risen to about 1 bar. These clouds of ammonia ice presumably rise to form the bright, white storms that are occasionally seen above the global haze. Water clouds may form much lower in the atmosphere,

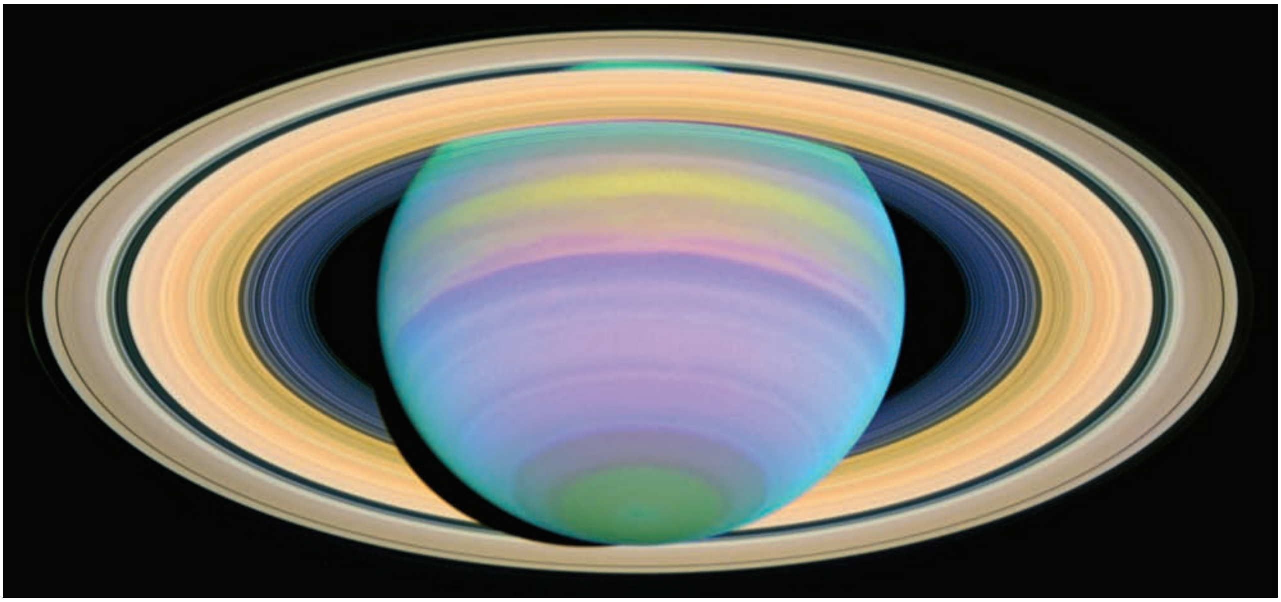


Fig. 10.3 Saturn at ultraviolet wavelengths The southern hemisphere of Saturn reaches its maximum 27-degree tilt toward Earth, revealing the planet's banded clouds as well as its rings. Particles in Saturn's atmosphere and rings reflect different wavelengths of radiation in different ways, as indicated in this ultraviolet image that reveals smaller particles that do not absorb or scatter radiation at the longer visible or infrared wavelengths. Saturn's counter-flowing east–west winds have aligned the cloud particles, believed to be ammonia ice crystals, within fixed latitude bands. The rings are made up of particles of water-ice. This view of Saturn was recorded from the *Hubble Space Telescope* on 7 March 2003. (Courtesy of NASA/E. Karkoschka, U. Arizona.)

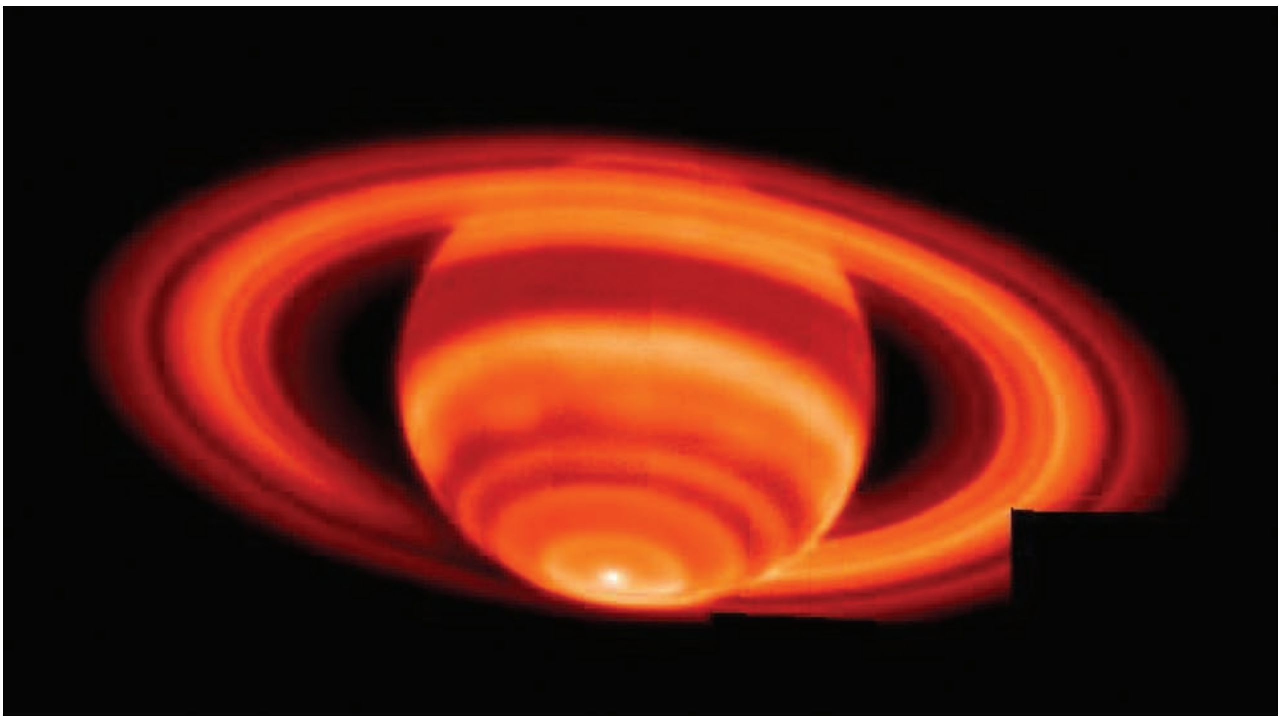


Fig. 10.4 Saturn at infrared wavelengths This image was taken at infrared wavelengths that are sensitive to temperatures in Saturn's upper troposphere. The prominent "hotspot" at the bottom of the image is located at the planet's south pole, which is in summer, but at the temperature of only 91 kelvin it is still freezing cold. Warm regions at 88 to 89 kelvin are found above 70 degrees southern latitude. Ring particles are not at uniform temperature everywhere in their orbit around Saturn. They are cold just after having cooled down in Saturn's shadow (*lower left*). As they orbit Saturn, the particles increase in temperature to a maximum just before passing behind Saturn again in shadow. This image is a mosaic of 35 individual exposures made at the Keck I Observatory in Mauna Kea, Hawaii, on 4 February 2004; the small missing section is due to incomplete mosaic coverage during the observing sequence. (Courtesy of NASA/JPL.)

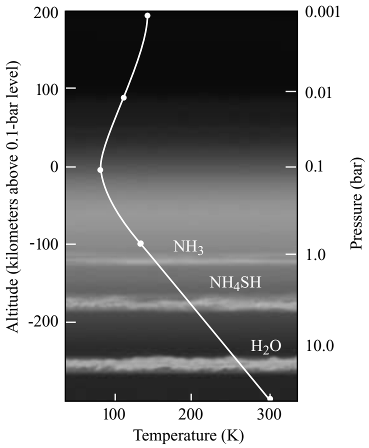


Fig. 10.5 Temperature and pressure at Saturn's cloud levels The fading radio signals when the *Voyager 1* and 2 spacecraft passed behind Saturn in 1980 and 1981, respectively, revealed the temperatures (*bottom axis*) and pressures (*right axis*) in its upper atmosphere. The temperature reaches a minimum of about 80 kelvin at a level called the tropopause where the atmospheric pressure is 0.1 bars, or 100 millibars. By way of comparison, the pressure of the Earth's atmosphere at sea level is 1.0 bar. The altitudes (*left axis*) are relative to the 0.1 bar level, and the dots are spaced to indicate tenfold changes in pressure. Solar radiation causes the temperature to increase with height just above the tropospause. At lower levels, the temperature and pressure increase systematically with depth. Three possible cloud layers of ammonia (NH_3), ammonium hydrosulfide (NH_4SH) and water-ice (H_2O) are shown. The altitudes of the predicted cloud layers are based on an equilibrium gaseous mixture that is of solar composition. An increase in abundance of a condensable gas by a factor of 3 would lower the altitude of the cloud base by about 10 kilometers.

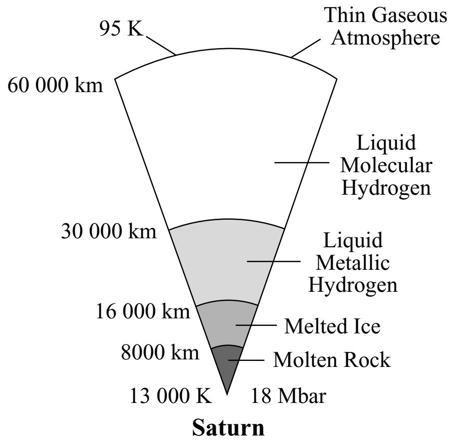


Fig. 10.6 Inside Saturn Giant Saturn has a thin gaseous atmosphere covering a vast global ocean of liquid hydrogen. At the enormous pressures within Saturn's interior, the abundant hydrogen is compressed into an outer shell of liquid molecular hydrogen and an inner shell of fluid metallic hydrogen. The giant planet may have a relatively small melted rock–ice core.

where the pressure rises to almost 10 bars, but no one has ever seen them.

10.3 Beneath Saturn's clouds

The internal constitution of Saturn

Saturn's low mass density indicates that the lightest element, hydrogen, is the main ingredient inside the planet, just as it is for Jupiter and the Sun. The lightweight material, just 68.7 percent as dense as water, is hurled outward in its equatorial regions by the planet's rapid 10.6562-hour rotation, making Saturn the most oblate planet in the solar system. Its equatorial bulge amounts to about 10 percent of the radius, and is about as big in extent as the Earth. Or, as some view it, the polar regions of Saturn are squashed and flattened by this amount.

The oblong shape of Saturn can be seen with a small telescope, and measured precisely from its satellite orbits and ring positions, as well as by the trajectories of the passing *Voyager 1* and *2* spacecraft and the orbiting *Cassini* spacecraft. When these measurements are combined with Saturn's known mass, volume and rotation rate, scientists can obtain information about its internal distribution of mass.

The model the experts come up with is just a scaled-down version of Jupiter, with a small, dense core of melted ice and molten rock surrounded by a vast globe of liquid hydrogen and topped by a thin gaseous atmosphere (Fig. 10.6). Like Jupiter, giant Saturn is not a solid world, and is essentially a large, round drop of liquid.

Deep down inside, the liquid hydrogen is compressed to such high pressures that it conducts electricity like a metal. But since Saturn's mass is less than a third of the mass of Jupiter, and only slightly smaller, the internal pressure at a given depth is less, and the liquid hydrogen turns into a metal further down in the ringed planet. Saturn therefore has a smaller shell of liquid metallic hydrogen.

Also like Jupiter, the giant planet Saturn has a magnetic field generated by rotationally driven electric currents in the planet's liquid metallic shell. But since Saturn has a thinner shell, the strength of its magnetic field is about one-twentieth of Jupiter's, despite the fact that both planets rotate with about the same period. Auroras are produced in Saturn's polar regions when charged particles spiral down along the magnetic field lines and collide with gases in the atmosphere.

One unusual characteristic of Saturn's magnetism is that its magnetic axis is almost precisely aligned with its

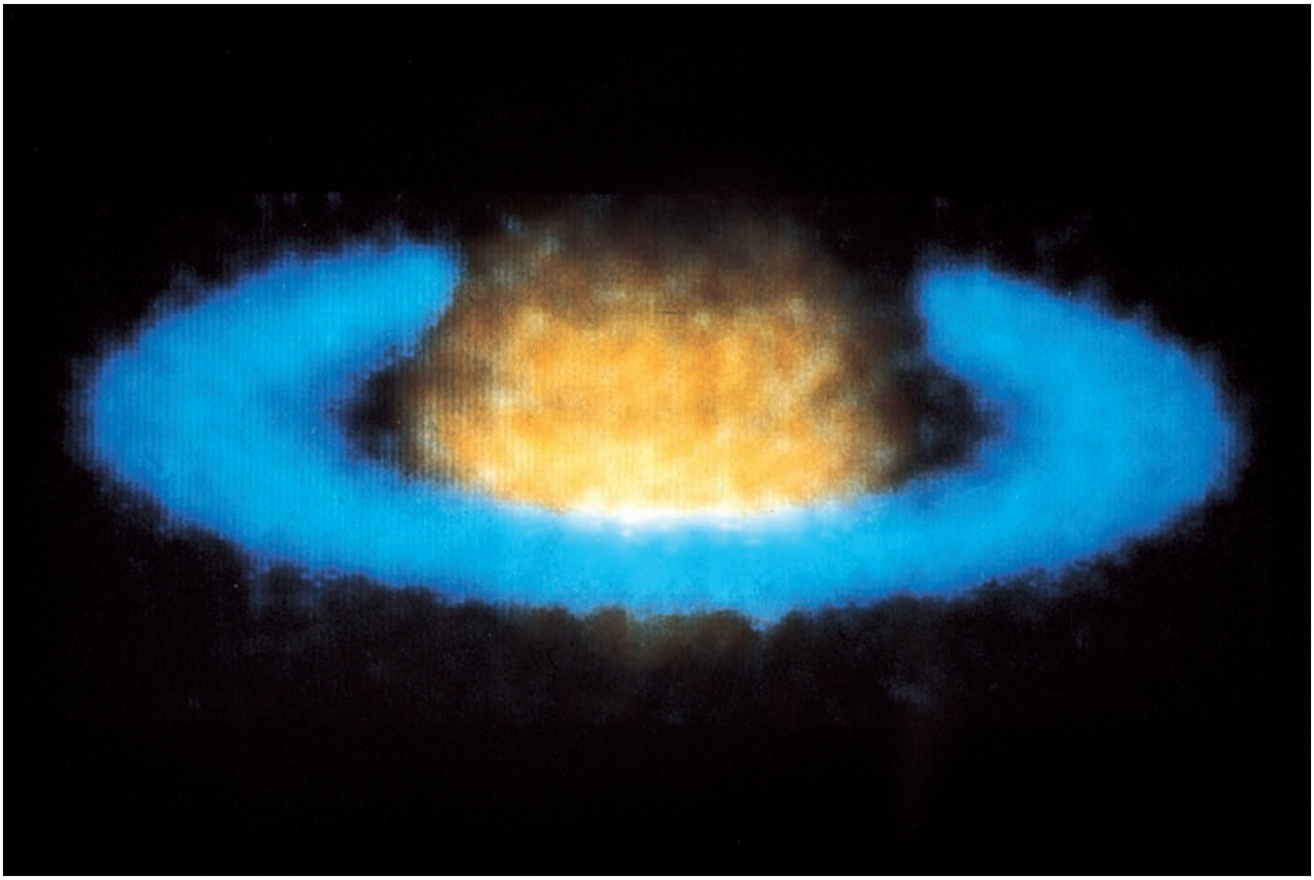


Fig. 10.7 Saturn and its rings at infrared wavelengths Because Saturn's rings are made of water-ice, they reflect relatively large amounts of sunlight at an infrared wavelength of 3.8 micrometers (*blue*). Methane in Saturn's outer atmosphere absorbs radiation at this wavelength, but the incandescent globe has its own internal source of heat that makes it shine brightly at the longer infrared wavelength of 4.8 micrometers (*orange*). The heat welling up from within Saturn is partly due to helium raining inside of the planet. (Courtesy of David Allen, Anglo-Australian Telescope Board © 1983.)

rotation axis. No other planet has such an alignment of the two axes, and it is difficult to explain how the magnetic field can be maintained in such a way.

Interior heat and helium rain

Precise measurements from the *Voyager 1* and *2* spacecraft indicate that Saturn is radiating 1.79 times more energy in visible and infrared light than it absorbs from incoming sunlight. This excess energy must be coming from within the planet. It implies that Saturn, like Jupiter, is an incandescent globe with an internal source of heat (Fig. 10.7).

Both Jupiter and Saturn radiate almost twice as much energy as they receive from the Sun, but the dominant source of internal heat is different for the two giant planets. Jupiter's internal heat is primarily primordial heat liberated during the gravitational collapse when it was formed, and Saturn must have also started out hot inside as the result of its similar formation. But being somewhat smaller and less massive than Jupiter, the planet Saturn was not as hot in its beginning and has had time to cool. As a result, Saturn lost most of its primordial heat, and there must be another source for most of its internal heat.

Saturn's excess heat is generated by the precipitation of helium into its metallic hydrogen core. The heavier helium separates from the lighter hydrogen and drops toward the center, somewhat like the heavier ingredients of a salad dressing that hasn't been shaken for a while. Small helium droplets form where it is cool enough, precipitate or rain down, like water vapor condensing out of the Earth's atmosphere, and then dissolve at hotter deeper levels. As the helium at a higher level drizzles down through the surrounding hydrogen, the helium converts some of its energy to heat. In much the same way, raindrops on Earth become slightly warmer when they fall and strike the ground; their energy of motion – acquired from gravity – is converted to heat.

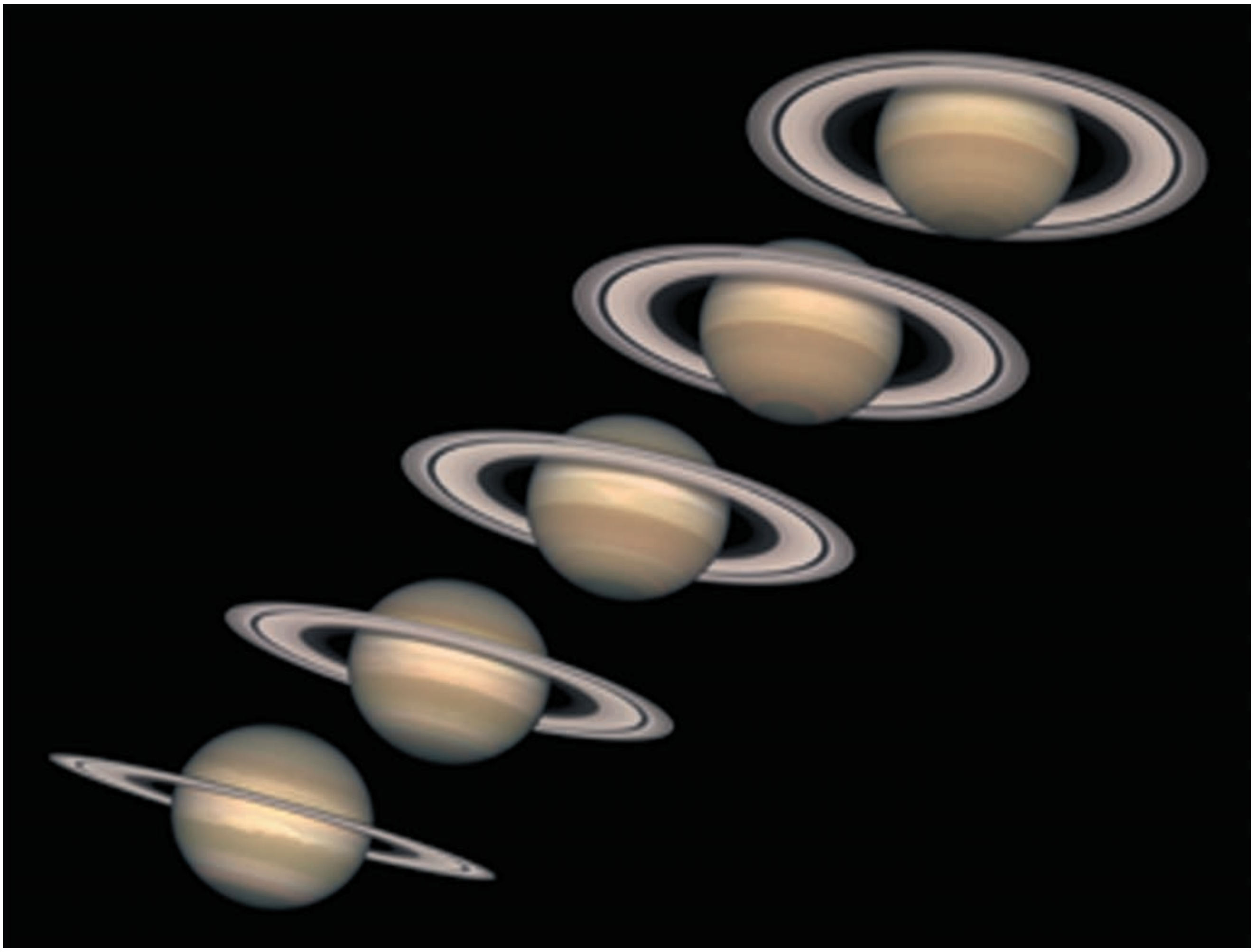


Fig. 10.8 Saturn's rings open up These images were taken from the *Hubble Space Telescope* during a four-year period, from 1996 to 2000 (*left to right*), as Saturn moved along one-seventh of its 29-year journey around the Sun. As viewed from near the Earth, Saturn's rings open up from just past edge-on to nearly fully open as it moves through its seasons, from autumn towards winter in its northern hemisphere. (Courtesy of NASA/STScI.)

The helium-rain theory has apparently been confirmed by *Voyager* measurements of a lower abundance of helium in the outer atmosphere of Saturn than in Jupiter or the Sun. The number of helium molecules in Saturn is only 3 percent (and hydrogen 97 percent), while the number is 16 percent in the Sun and 13.6 percent for Jupiter. Since Jupiter is just slightly depleted in helium when compared to the Sun, helium rain is also probably operating inside this giant planet, but in more modest amounts because of Jupiter's greater mass and internal temperature.

At the higher temperatures and pressures found in Jupiter's interior, most of the liquid helium dissolves in liquid hydrogen, like cooking ingredients which dissolve more easily in hot liquids than cold ones. Inside Saturn, which has a lower internal temperature, some of the helium forms droplets instead of dissolving.

10.4 The remarkable rings of Saturn

Billions of whirling particles of water ice

The austerely beautiful rings of Saturn are so large and bright that we can see them with a small telescope. Because the glittering rings are tipped with respect to the ecliptic, the plane of the Earth's orbit about the Sun, they change their shape when viewed from the Earth (Fig. 10.8). The rings are successively seen edge on, when they can briefly vanish from sight in a small telescope, from below, when they are wide open, edge-on again and then from above. The complete cycle requires 29.4475 Earth years, the orbital period of Saturn, so the rings nearly vanish from sight every 15 years or so.

The three main rings of Saturn have been observed for centuries. They are the outer A ring and the central

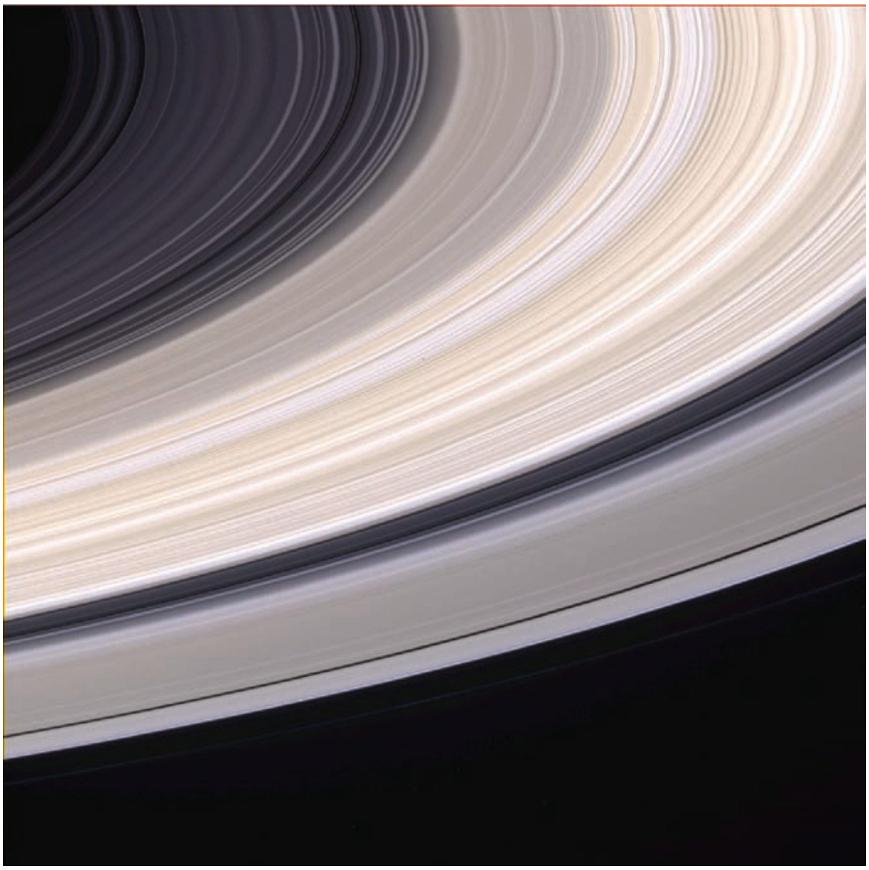


Fig. 10.9 Saturn's rings of ice The narrow-angle camera aboard the *Cassini* spacecraft took this image from beneath Saturn's ring plane on 21 June 2004. The brightest part of the ring system, extending from the upper right to the lower left, is the central B ring. It is separated from the outermost A ring by the wide, dark Cassini Division, discovered in 1675 by the Italian-born French astronomer Giovanni (Gian) Domenico (Jean Dominique) Cassini (1625–1712). Below the B ring, closer to the planet, is the C ring. All three rings are composed of innumerable particles of water-ice. The different shades of the rings are attributed to different amounts of contamination by other materials such as rock or carbon compounds. When viewed close up, the broad icy rings break up into thousands of individual wave-like ringlets. (Courtesy of NASA/JPL/SSI.)

B ring, separated by the dark Cassini Division, and an inner C, or crepe, ring that is more transparent than the other two. They remain suspended in space, unattached to Saturn, because they move around the planet at speeds that depend on their distance, opposing the pull of gravity.

The motions of Saturn's rings can be measured using spectral features in their reflected sunlight. When part of a ring moves toward or away from an observer, the spectral features are displaced in wavelength by an amount that depends on the velocity of motion. There is a shift toward shorter wavelengths when the motion is toward the observer, while motion away produces a shift toward longer wavelengths. Observations of this Doppler effect, by the American astronomer James Edward Keeler (1857–1900) in 1895, showed that the inner parts of the rings move around Saturn faster than the outer parts, all in accordance with Kepler's third law for small objects revolving about a massive, larger one. They orbit the planet with periods ranging from 5.8 hours for the inner edge of the C ring, to 14.3 hours for the outer edge of the more distant A ring. Since Saturn spins about its axis with a period of 10.6562 hours, the inner parts of the main rings orbit at a faster speed than the planet rotates, and the outer parts at a slower speed.

The difference in orbital motion between the inner and outer parts of the rings means that they are not a solid sheet of matter, for they would be torn apart by the differential motion. As demonstrated by James Clerk Maxwell (1831–1879) in 1867, the rings are instead made up of vast numbers of particles, each one in its own orbit around Saturn, like a tiny moon. Billions of ring particles revolve about the planet. They have been flattened and spread out to a thin, wide disk as the result of collisions between particles.

The rings of Saturn are flat, wide and incredibly thin. Measured from edge to edge, the three main rings span a total width of 62 200 kilometers, so they are a little wider than the planet's mean radius, at 60 228 kilometers (Fig. 10.9; Table 10.2). When observed edge on, from on or near the Earth, the rings practically disappear from view (Fig. 10.10). They look about a kilometer thick, but this thickness has been attributed to warping, ripples, embedded satellites and a thin, inclined outer ring. When instruments on *Voyager 2* monitored starlight passing through the rings, they found that the ring edges extend only about 10 meters from top to bottom. If a sheet of paper represents the thickness of Saturn's rings, then a scale model would be two kilometers across.

When sunlight hit the rings exactly edge-on, instruments on the *Cassini* spacecraft showed that there is a lot

Table 10.2 Saturn's rings[a]

Ring	Width (km)	Closest distance (km)	Distance range (R_S)	Particle size (m)	Optical depth	Mass (kg)
D	7540	66 970	1.11 to 1.235	$<10^{-6}$	0.0001	
C	17 490	74 510	1.235 to 1.525	0.01 to 3.0	0.05 to 0.35	1×10^{17}
B	25 580	92 000	1.525 to 1.949	0.01 to 5.0	0.4 to 2.5	2×10^{19}
A	14 610	122 170	1.949 to 2.025	0.01 to 7.5	0.05 to 0.15	4×10^{17}
F	50	14 180	2.324	10^{-7} to 10^{-5}	0.1	
G	500 to 3000	170 180	2.82	3×10^{-8}	2×10^{-6}	
E	302 000	181 000	3 to 8	1×10^{-6}	1.5×10^{-5}	7×10^{8}

[a] The ring widths and closest distances are given in kilometers (km), the particle size in meters (m), and the mass in kilograms (kg). The distance range is given in units of Saturn's apparent equatorial radius, $R_S = 60\,330$ kilometers. At the one-bar pressure level, the equatorial radius is 60 268 kilometers.

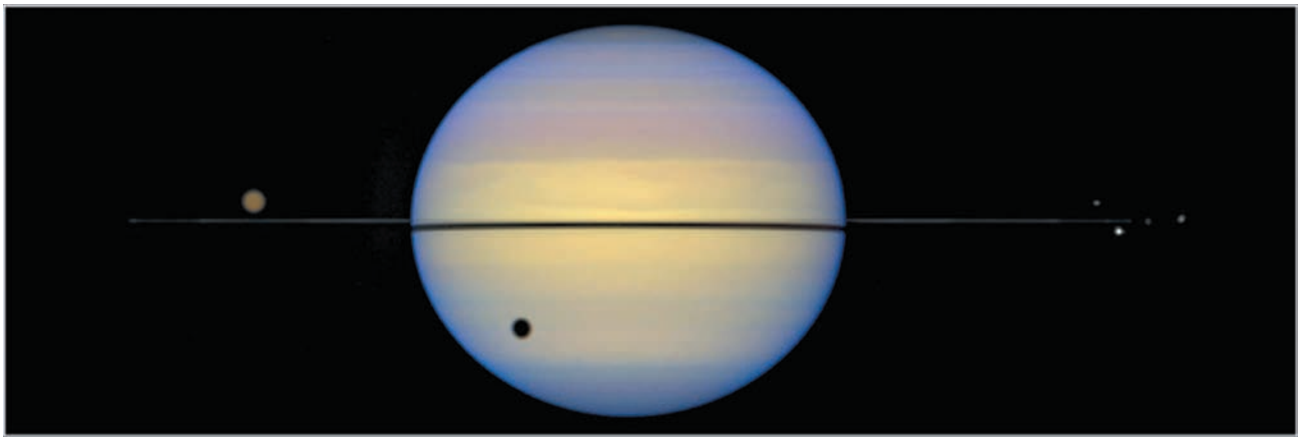


Fig. 10.10 Edge-on view of Saturn's rings When the Earth is in the plane of Saturn's rings, an observer on the Earth views the rings edge on. Because the rings are so thin, they are then barely visible. Saturn's largest satellite, Titan, is seen just above the rings (*left*); it is enveloped in a dark brown haze and casts a dark shadow on Saturn's clouds. Four other moons are clustered near the other edge of Saturn's rings (*right*), appearing bright white because their surfaces are covered with water-ice. From left to right, these icy satellites are named Mimas, Tethys, Janus and Enceladus. This image was taken on 6 August 1995 from the *Hubble Space Telescope*. (Courtesy of NASA/STScI.)

of vertical relief above and below the paper-thin rings. Rippling corrugations undulate though the innermost D ring and the neighboring C ring, even into the B ring. Bumps and ridges pile up to unexpected heights. One ridge of icy ring particles loomed as high as 4 kilometers, apparently pulled up by a moon as it traveled through the plane of the rings. So the rings aren't perfectly flat, and depart from thinness in places.

What are the ring particles made out of? At visible wavelengths, the rings are bright and reflective, but at infrared wavelengths they are dark and less reflective. This suggests that the particles are cold and made of ice. In fact, they are composed largely, and almost exclusively, of water ice.

Detailed Earth-based infrared spectroscopy of the main rings in the 1970s showed that incident sunlight is absorbed by water ice at the surfaces of the particles. Subsequent spectral investigations indicated that the frozen water is exceptionally clean and pure, with negligible amounts of dust or rock. They are also poor absorbers and emitters of microwaves, which implies that more than 99 percent of the mass of the rings is water ice, and that less than 1 percent consists of dirty contaminants.

The total mass of the prominent A, B and C rings is about equal to that of Saturn's medium-sized satellite Mimas, which has a mean radius of 198 kilometers and weighs in at 3.749×10^{19} kilograms. Such a mass is consistent with particles composed of water ice. To check that, just multiply the mass density of water, at 1000 kilograms per cubic meter, by the total volume of the main rings – 10 meters thick, 60 000 kilometers wide, and a

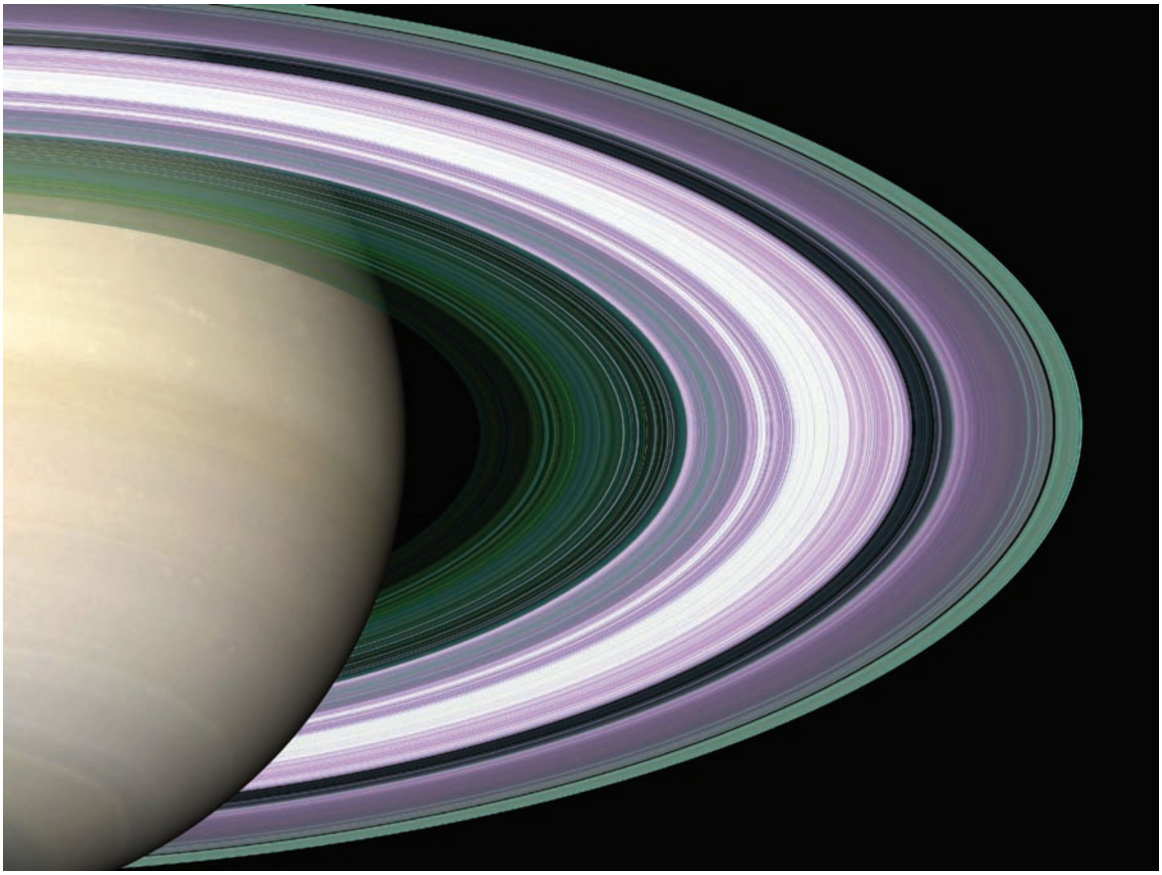


Fig. 10.11 Particles in Saturn's rings When the *Cassini* spacecraft passed behind the rings of Saturn, it sent three simultaneous radio signals, at 0.94, 3.6 and 13 centimeter wavelength, through the rings to Earth. The observed changes of each signal as the spacecraft moved behind the rings provided a profile of the distribution of ring material as a function of distance from Saturn, or an optical-depth profile. The image shown here was constructed from these profiles, depicting the observed ring structure at about 10 kilometers in resolution. Color is used to present information about the presence or absence of small ring particles in different regions based on the measured effects of the three radio signals. Purple color indicates regions where there is a lack of particles of size less than 5 centimeters. Green and blue shades indicate regions where there are particles smaller than 5 centimeters and 1 centimeter, respectively. The saturated broad white band is the densest region of the B ring, which blocked two of the three radio signals. From other evidence in the radio observations, all ring regions appear to be populated by a broad range of particle sizes that extend to several meters across. (Courtesy of NASA/JPL.)

circumference of about 600 000 kilometers. Since the particles are not jammed tightly together, and probably separated by 5 to 10 times their size, the resulting mass has to be diluted by a corresponding factor.

Typical chunks of ice in the main rings vary in size from hailstones to fist-sized chunks of ice; some are as small as snowflakes, and a few of the icebergs are as large as a house. In other words, the ring particles range in size from a hundredth of a meter to ten meters across (Table 10.2). There are more and more particles of smaller and smaller size within this range, so the main rings consist primarily of the smaller particles about 0.01 meters in size. Though far less numerous, the larger particles greater than 1 meter across contain most of the ring mass.

The ring particles are too small for spacecraft cameras to see individually, but scientists can infer their size from radio measurements. Since the rings are very reflective to ground-based radar transmissions, we know that their particles are comparable to, or larger than, the radar wavelength of about 0.1 meters. The particle-size distribution has been determined from the way the rings blocked the radio signals from *Cassini* when the spacecraft passed behind the rings (Fig. 10.11). This method showed that there are remarkably few particles either larger than 5 to 10 meters in size or smaller than 0.01 meters. Within these bounds, the number of particles in the main rings decreases with increasing size, in proportion to the inverse square of their radius.

However, four additional rings, designated the D, E, F and G rings (Fig. 10.12), consist of much smaller, microscopic ice crystals. These rings, discovered using ground-based or spacecraft observations, are all very

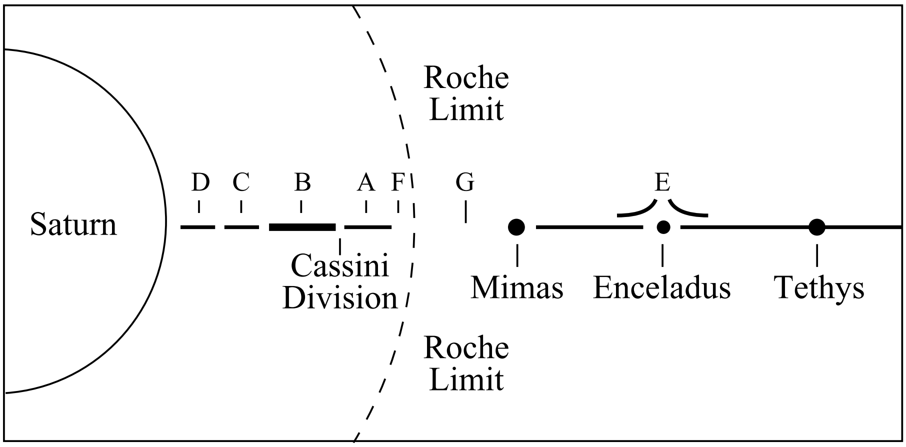


Fig. 10.12 Cross-section of rings and satellites All of Saturn's main rings lie inside the Roche limit (*dashed curve*) within which the planet's gravity will tear a large satellite apart. The A and B rings have been observed for centuries. The more tenuous C ring was discovered in the 19th century, and definite observations of the transparent D ring awaited the arrival of the *Voyager 1* spacecraft on 12 November 1980. The icy satellite Enceladus feeds the tenuous E ring, also revealed from *Voyager 1*, as well as from the *Cassini* spacecraft. For clarity, the thickness of the rings has been exaggerated.

diffuse, tenuous and nearly transparent. The way that their particles scatter light indicates that they are the smallest of all, roughly a micron in size – a micron is a millionth of a meter in size, or 10^{-6} meters.

Lying between the C ring and the planet is the D ring. Although terrestrial observers had reported a faint ring between C and the globe as early as 1969, these reports remained controversial until the *Voyager* spacecraft definitely verified the existence of the D ring. It is so tenuous and transparent that it is probably impossible to see from the Earth using the best telescopes. According to one hypothesis, splintered chips from colliding ice particles drift down into Saturn's atmosphere and form the D ring.

Outside the traditional system lies the huge, sparse E ring, a broad tenuous band of small particles that is five times the combined breadth of rings A, B and C, and roughly centered on the orbit of Enceladus. Discovered with ground-based telescopes in 1966, the E ring becomes visible when the ring system is viewed approximately edge-on. It is composed almost exclusively of small grains just one micron in size.

Because they have relatively short lifetimes of several thousand years, these tiny particles must be continually replenished if the E ring is a permanent feature. As discussed in greater detail later in this chapter, watery eruptions from ice jets and geysers on Enceladus feed small bits of ice into Saturn's E ring. Once lofted into space, the pressure of sunlight, the gravitational tugs of Saturn, and possibly electromagnetic effects spread the particles out.

Initially discovered by instruments on the *Pioneer 11* spacecraft and verified in *Voyager* images, the tenuous G ring lies between the A and E rings. The G ring consists of similar micron-sized particles to the E ring, and may be renewed from small moons embedded within it.

The *Pioneer 11* instruments also discovered the incredibly narrow F ring, that lies just outside the A ring, by its absorption of energetic particles; while images from the *Voyager* spacecraft showed the F ring in great detail, demonstrating that its width varies from a few to tens of kilometers. Moreover, it is not just a single ring; *Voyager 1* spotted a contorted tangle of narrow strands that had smoothed out by the time *Voyager 2* arrived about 9 months later. Because the F ring particles are brighter when backlit by the Sun, and fainter in reflected sunlight, we know that the particles are also micron-sized, much smaller than snowflakes and comparable in size to the dust in your room.

But how can this ring retain such narrow boundaries? In the absence of other forces, collisions between ring particles should spread them out, causing the particles to fall inward toward Saturn and expand outward from it, thus creating a broader and more diffuse ring. Two tiny moons, named Pandora and Prometheus, flank the F ring and confine it between them, thereby keeping the particles of the F ring from straying beyond the ring's narrow confines (Fig. 10.13).

These shepherd satellites, discovered by the *Voyager* spacecraft, chase each other around the narrow F ring and keep it from spreading, as though they were two gravitational sheepdogs herding sheep into a narrow path. Each shepherd tends one edge of the ring. The moon outside the ring moves more slowly than the ring particles, which in turn are outpaced by the fast inner moon. The gravity of

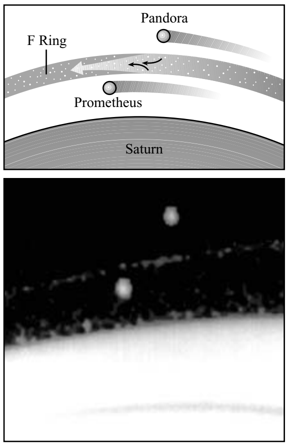


Fig. 10.13 Saturn's shepherd satellites Two shepherd satellites confine Saturn's narrow F ring. The outer shepherd gravitationally deflects ring particles inward, and the inner shepherd deflects ring particles outward. (Courtesy of NASA/JPL.)

the faster-moving inside satellite pulls the inner F ring particles forward as it passes, causing them to accelerate and spiral outward. The slower-moving outer shepherd exerts a net backward force on the outer ring particles, causing them to move inward. The result is a very narrow ring. Such confining moons were originally proposed to account for the narrow rings surrounding Uranus, constraining their edges from the otherwise inevitable spreading, and such a pair of satellites was eventually found astride one of its rings.

The shepherd satellites that flank the F ring, and possibly one or more embedded satellites as well, gravitationally interact with the ring material and distort its normal, circular ring shape, producing temporary kinks and twists in the clumpy, braided ring. Instruments aboard the *Cassini* spacecraft have revealed details of the interaction, showing that Prometheus can draw out material when passing closest to the F ring. Because their orbits are slightly eccentric, these two satellites produce a varying perturbation of ring particles, perhaps accounting for the changing smooth and contorted appearance of the F ring. Through similar gravitational interactions, small moons can produce ripples and waves on the surface of Saturn's main rings, and clear gaps within them.

Ringlets, waves, gaps and spokes

From a distance, the principal rings of Saturn look like smooth, continuous structures. Up close, however, from the views provided by the *Voyager 1* and *2* and *Cassini* spacecraft, the icy material is marshaled into thousands of individual ringlets (Fig. 10.14). Some of the ringlets are perfectly circular, others are oval-shaped and a few seem

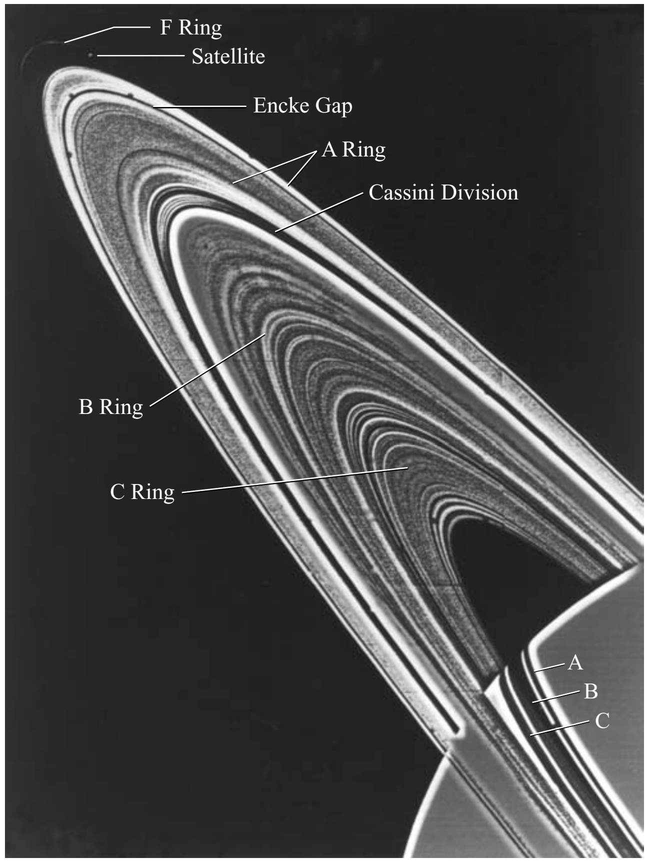


Fig. 10.14 Ringlets When viewed with high resolution, approximately 100 concentric features are seen within Saturn's rings, including some in the Cassini Division. The ring system would probably separate into countless ringlets if we could detect fine enough detail. A small satellite, discovered by *Voyager 1*, is seen (*upper left*) just outside the narrow F ring. The *Voyager 1* spacecraft took this mosaic of Saturn's rings on 6 November 1980. (Courtesy of NASA/JPL.)

to spiral in towards the planet like the grooves on an old-fashioned record. In some places, the flat plane of the rings is slightly corrugated, and ringlets are seen at the crests and dips of the corrugations, like ripples running across the surface of a pond.

An outside hand is at work sculpting at least some of the intricate ring structures through the force of gravity. The combined gravitational pull of Saturn and the accumulated pull of nearby moons can redistribute the ring particles, concentrating them into many of the observed shapes. Although small nearby moons have only a weak gravitational pull on the particles in the rings, the pull becomes pronounced when the particle orbital period and the satellite orbital period are exact fractions or whole-number ratios, such as $\frac{1}{2}$ for the 2 : 1 resonance, and the gravitational perturbations are repeated over and over again in the same resonant location. The interplay of this

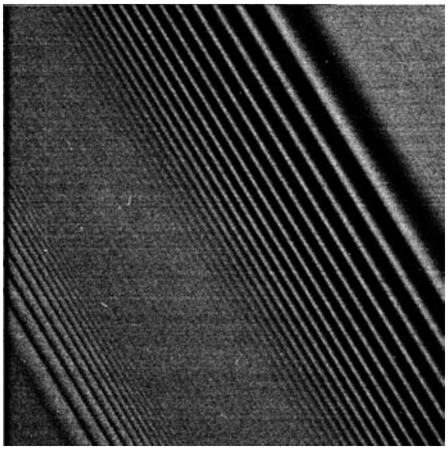


Fig. 10.15 Two waves in Saturn's rings Gaps and wave-like concentrations in ring particles are due to the gravitational influence of Saturn's moons. A small, nearby moon orbiting at varying distance from Saturn's rings is thought to produce waves of density, causing the ring particles to bunch together and disperse like the crests and troughs of ocean waves (*bottom left*). The gravity of a nearby moon can also produce a vertical wave in the ring particles, known as a bending wave (*upper right*). This image, which spans about 220 kilometers, was taken form the *Cassini* spacecraft on 29 October 2004. (Courtesy of NASA/JPL/SSI.)

effect and Saturn's inward gravitational pull can repel and attract the ring particles, pushing and pulling them into localized alternating high and low concentrations such as ringlets.

The additive gravitational perturbations between a nearby moon and the ring particles can also make waves of density propagate radially through the rings (Fig. 10.15). The particles tend to congregate at the crests and troughs of these density waves, like automobiles in traffic intersections or a crowd starting a "wave" in a stadium. Similar spiral density waves on a vastly greater scale are thought to create the stellar arms of spiral galaxies.

Resonance can also confine ring edges and clear gaps. Particles straying into a gap at a resonant location are removed by repetitive interaction with a particular moon. The ring particles at the outer edge of the B ring and the inner edge of the Cassini Division, for example, are traveling almost twice as fast as Saturn's largest inner satellite, Mimas, with a period one-half as large and occupying the 2 : 1 resonance with this moon. The razor-sharp outer edge and scalloped hem of the A ring similarly result from a 7 : 6 resonance with the external co-orbital satellites Janus and Epimetheus. In addition, many low-density regions in Saturn's A ring are located at positions resonant with small moons that lie just exterior to the ring.

Tiny satellites embedded within the rings can also sweep out a tidy gap with neat edges. The small moon Pan, embedded within Saturn's A ring, apparently plows its way through Encke's gap, keeping it open and creating wavy radial oscillations around the inner and outer edges of the gap.

But simple interactions with known moons have not been completely successful in accounting for all of the intricate detail found in Saturn's rings. The apparent gaps in the system are not completely empty. The Cassini Division, for example, contains perhaps 100 ringlets (Fig. 10.16), with particles just as large as those in the neighboring ring. Some gaps do not even occur at known resonant positions or contain detected moons embedded within them. Unseen moons might influence the clumping and removal of material in these locations.

Backlit images also reveal several faint rings of microscopic particles within the Cassini Division, overlying the orbits of small inner moons such as Janus and Epimetheus and Pallene, and ring arcs extending from other small moons along their orbits, like Anthe and Methone.

Perhaps the most bizarre *Voyager* discovery was the long, dark streaks, dubbed spokes, which stretch radially across the rings, keeping their shape like the spokes of a wheel. These ephemeral features are short-lived, but regenerated frequently. They are found near the densest part of the B ring that co-rotates with the planet at a period of 10.6562 hours. But the inner and outer parts of Saturn's dark spokes also whirl around the planet with this period, at constant speed in apparent violation of Kepler's third law and Newton's theory of gravity. If the spokes consisted of dark particles embedded in the rings, the particles would move with speeds that decrease with increasing distance from Saturn, and the spokes would quickly stretch out and disappear.

The exact mechanism for generating and sustaining the mysterious spokes remains obscure, but according to one hypothesis, the small dust particles may become charged, perhaps as the result of collisions with energetic electrons. Electromagnetic forces then raise or levitate the tiny, charged particles off the larger ring bodies, and the spokes are swept around Saturn by its rotating magnetic field. It sounds bizarre, but subtle forces are required to overcome gravity.

Why do planets have rings?

One might expect the particles of a ring to have accumulated long ago into larger satellites. But the interesting feature of rings – and a clue to their origin – is that they do not coexist with large moons. Planetary rings are usually closer to the planets than their large satellites.

The rings are normally confined to an inner zone where the planet's tidal forces would stretch a large satellite until

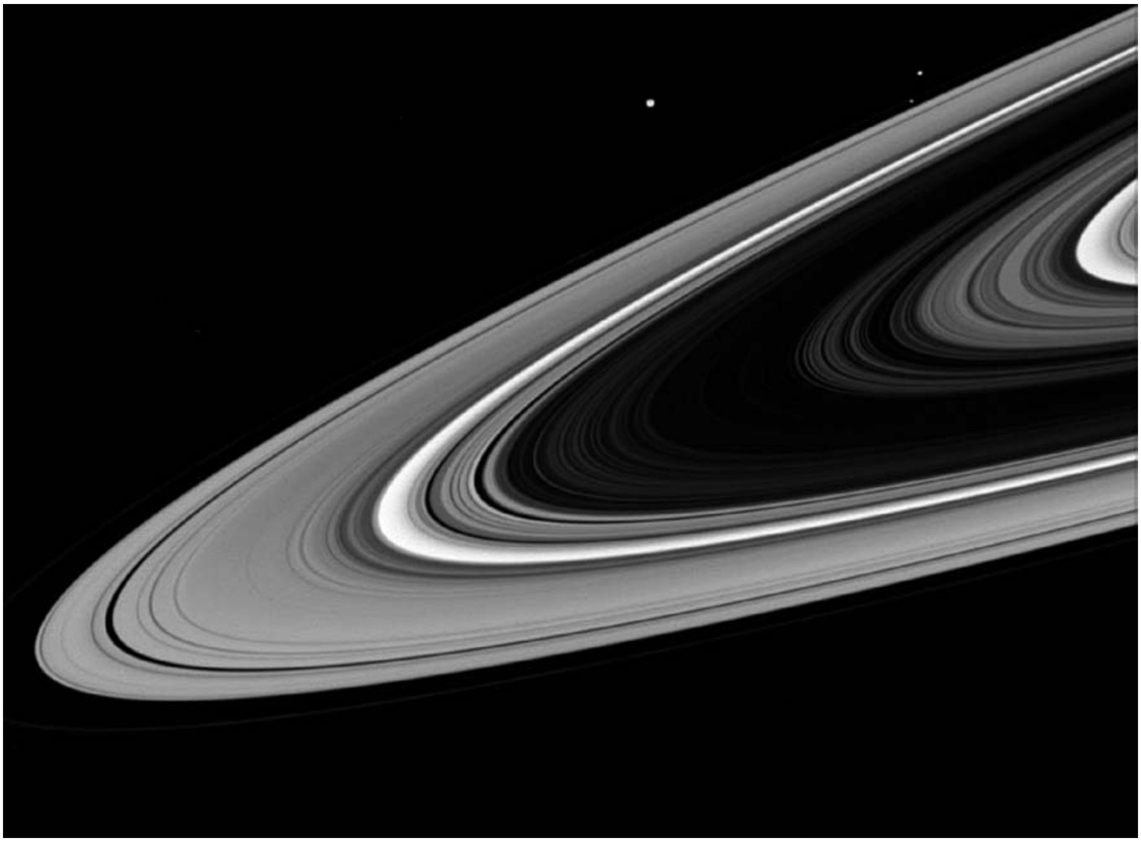


Fig. 10.16 Beneath the rings of Saturn When the *Voyager 1* spacecraft dove beneath Saturn's rings, it could view sunlight transmitted through the rings, presenting a reversed image of the sunlit side. Both the C ring and Cassini Division appear bright because they are sparsely populated with small particles that efficiently scatter light in the forward direction, whereas the B ring appears dark because its densely packed particles absorb all the incident sunlight. This perspective is not available from Earth, where we always see the sunlit side of the rings. Two bright, icy satellites are seen at the upper center and upper left. (Courtesy of NASA/JPL.)

it fractured and split, while also preventing small bodies from coalescing to form a larger moon. The outer radius of this zone in which rings are found is called the Roche limit after the French mathematician Eduoard A. Roche (1820–1883), who described it in 1848 (Fig. 10.17).

For a rigid satellite with the same mass density as its planet, the Roche limit is 1.26 times the planet's radius (Focus 10.1). The Roche limit for a solid body is 1.38 times that radius, and Roche's initial calculation, for fluid objects, was 2.446 times the planetary radius. Anywhere inside this distance a large satellite can no longer remain intact, but instead gets torn apart by planetary tides. Nevertheless, because of their material strength and great internal cohesion, small moons less than 100 kilometers across can exist inside the Roche limit without being tidally disrupted, just as the ring particles can.

For a satellite with no internal strength and whose density is the same as the planet, the Roche limit is 2.446 times the planetary radius, or about 175 000 kilometers for Jupiter, 147 000 kilometers for Saturn, 62 000 kilometers for Uranus, and 59 000 kilometers for Neptune. Jupiter's

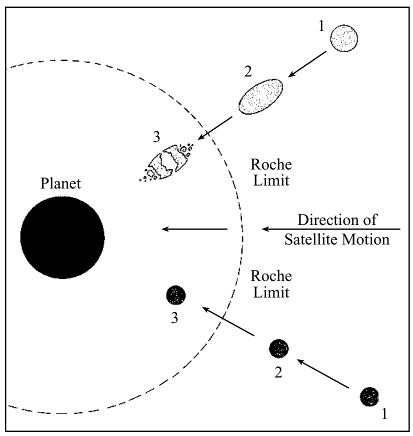


Fig. 10.17 The Roche limit A large satellite (*top*) that moves well within a planet's Roche limit (*dashed curve*) will be torn apart by the tidal force of the planet's gravity. The side of the satellite closer to the planet feels a stronger gravitational pull than the side farther away, and this difference works against the self-gravitation that holds the body together. A small solid satellite (*bottom*) can resist tidal disruption because it has significant internal cohesion in addition to self-gravitation.

Focus 10.1 The Roche limit

To visualize the significance of the Roche limit, consider two particles of mass m, separated by a distance R, and located at a distance D from a planet of mass M_P. The gravitational pull of the planet on the particle closest to it will be greater than the pull on the more distant particle. If the difference in pull on the near and far particles, the tidal force, exceeds the mutual gravitational attraction between the two particles, they cannot remain close to each other and will disperse. The outcome of the tug-of-war between the tidal force and the mutual attraction is primarily decided by the particles' distance from the planet. At distances less than the Roche limit, D_{Roche}, particles are pulled apart, and this prevents the accumulation of larger moons. The tidal force will also tear apart any large moon-like object that ventures within the Roche limit.

The gravitational force F_P of a planet of mass M_P on a smaller mass m, whose center is located at a distance D from the center of the planet, is

$$F_P = \frac{G M_P m}{D^2}$$

where G is the Newtonian constant of gravitation. The planet will pull harder on the side of the object that is closer to it and less hard on the side that is further away. The difference ΔF between the force felt by one side and the center of the mass m is

$$\Delta F = \frac{G M_P m}{2}\left[\frac{1}{(D+R_m)^2} - \frac{1}{(D-R_m)^2}\right] = \frac{2 G M_P m}{D^3} R_m$$

where R_m is the radius of the smaller object. If it approaches the planet, D becomes smaller and this tidal force will increase, eventually pulling the object apart at a critical distance D_{Roche} from the center of the planet.

The gravitational binding force F_B, which attracts the extreme sides of the object and holds it together, is $Gm/{R_m}^2$ per unit mass, or for the total mass

$$F_B = \frac{G m^2}{R_m^2}$$

The Roche limit is reached when the tidal disruptive force, ΔF, equals the binding force, F_B, and when we set these two expressions equal and collect terms we obtain

$$D_{Roche} = \left(\frac{2 M_P}{m}\right)^{1/3} R_m$$

This result is expressed in terms of R_m, the radius of the small object, but by using the mass densities ρ_P and ρ_m, with $M_P = 4\pi \rho_P {R_P}^3/3$ and $m = 4\pi \rho {R_m}^3/3$, we obtain the Roche limit in terms of the planet radius, R_P:

$$D_{Roche} = \left(\frac{2\rho_P}{\rho_m}\right)^{1/3} R_P$$

which for a planet and smaller object of the same mass density becomes

$$D_{Roche} = 1.26\ R_P$$

The calculation by Roche used liquid objects whose shapes can distort continuously, and his result is

$$D_{Roche} = 2.446 \left(\frac{\rho_P}{\rho_m}\right)^{1/3} R_P \approx 2.446\ R_P$$

mere wisp of a ring, the icy snowballs of Saturn's rings, and the dark boulders in the narrow rings encircling Uranus and Neptune all lie within the Roche limit for the relevant planet. The Earth's Roche limit is 18 470 kilometers, and if our Moon ever ventured within this distance from the Earth's center, it would be pulled apart by tidal forces and our planet would have rings.

And where did Saturn's rings come from? There are two possible explanations for their origin. In the first explanation, the rings consist of material left over from Saturn's birth about 4.6 billion years ago. This hypothesis assumes that the rings and moons originated at the same time in a flattened disk of gas and dust with large, newborn Saturn at the center. According to the second explanation, a former moon or some other body moved too close to Saturn and was torn into shreds by the giant planet's tidal forces, making the rings. In this case, the rings could have formed after Saturn, its satellites and much of the rest of the solar system.

It has long been thought that Saturn's rings and satellites are both primordial leftovers of the planet formation process. Any disk material initially within the Roche limit soon after Saturn formed would have been prevented by the giant planet's tidal stresses from gathering or accreting into a large satellite. But outside the Roche limit, satellites could have coalesced from smaller bodies. Thus, any primordial, circumplanetary disk should develop into rings near the planet and exterior satellites, as we see today around Saturn. This could also help explain why only the giant planets have rings and a retinue of satellites, while the rocky terrestrial planets have no rings and either no satellite or just one or two of them. But there is a recent controversy over this long-standing explanation, centered on the fact that the planetary rings might be relatively young and therefore cannot be as ancient and enduring as the planets themselves.

Data obtained from the *Voyager* missions suggested that the shimmering rings of Saturn are ephemeral, created about 100 million years ago when a huge moon or comet came too close to the planet and shattered into pieces. The dazzling, sparkling brightness of Saturn's rings suggested such relatively young rings. They glisten with clean particles of pure water ice, unsullied by the constant pelting of cosmic dust. The rings could look much darker if they were very old, just as new-fallen snow becomes dirty over time.

The gravitational tugs of Saturn's moons on the rings will shorten the lives of the rings, providing another argument for their youth. When setting up density waves in the rings, nearby moons extract momentum from the ring particles, causing them to slowly spiral toward Saturn. To conserve momentum in the overall system, the moons gradually move away from the planet.

But subsequent data obtained from the *Cassini* spacecraft provided a different perspective, suggesting that some of the material in the rings is very long-lived, staying near the planet for billions of years. Ultraviolet light reflected off and passing through the water-ice particles indicated three times the mass assumed from the *Voyager* observations. This provides more raw materials for continuously recycling the ring material.

In this picture, small moons may be endlessly shattered into ring particles and gathered together, reforming. A larger ring mass indicates that the rings clump, rather than being uniformly distributed. There is a competition between the clumping of the ring particles due to gravitational interaction and collisions by micrometeorites, which shatter and disperse these clumps. The two processes can go on for several billion years.

So, it was once imagined that Saturn's resplendent rings formed with the planet about 4.6 billion years ago and the rings were next thought to be no older than 100 million years or they would not be there now. If we lived at the time of the dinosaurs, we might not have seen any rings around Saturn, and the rings we see today were once thought to be just temporary embellishments destined to disappear from sight in 100 million years or so. But now some of the very scientists who were once arguing for the rings' youth now say that they may have been created roughly 4.6 billion years ago when the solar system was formed, which means that Saturn may have always had rings.

10.5 Introduction to Saturn's moons

Saturn's medium-sized icy moons

Saturn has only one large satellite, Titan, comparable in size to Jupiter's four Galilean satellites, but it has an extensive family of smaller moons, including six mid-sized icy bodies that range from 98 to 765 kilometers in radius. In order of increasing orbital distance from Saturn, and also of increasing size, they are Mimas, Enceladus, Tethys, Dione, Rhea and Iapetus (Table 10.3). They were all discovered long ago – one by Christiaan Huygens (1629–1695), two by William Herschel (1738–1822) and four by Giovanni Domenico Cassini (1625–1712). Saturn's smaller retrograde satellite Phoebe wasn't discovered until 1899, by William H. Pickering (1858–1938).

The mean mass densities of these moons are often low, between 1100 and 1300 kilograms per cubic meter, which suggests that they are mainly composed of pure water ice. A few of them, such as Enceladus, Titan and Phoebe, have higher mass densities, indicating a composition of rock and ice. This is consistent with their highly reflective surfaces. With the exception of Iapetus and Phoebe, they all reflect more than 50 percent of the incident sunlight, and one of them, Enceladus, reflects almost 100 percent of the sunlight that strikes it. Surface water ice was also identified by infrared spectroscopy in the years prior to the *Voyager* missions.

The cold surfaces of all of Saturn's mid-sized icy satellites preserve ancient impact craters dating back to their formative years. Several of them also exhibit evidence for melting ice and surface activity after their formation.

There are several ways to group these objects. Mimas, Rhea and Iapetus, for example, have surfaces dominated by craters with no signs of internal activity. Or the moons can be paired by approximately the same size – Mimas and Enceladus, Tethys and Dione, and Rhea and Iapetus.

Dione provides a representative example of the icy cratered surfaces of these moons (Fig. 10.18). It has a mean mass density of 1476 kilograms per cubic meter, with perhaps enough rock in its makeup to produce internal heat from radioactivity. Most of the surface is heavily cratered, with varying numbers suggesting resurfacing by flows of erupting material.

Small moons of Saturn

Instruments on the *Voyager 1* and *2* and *Cassini* spacecraft have discovered a host of small moons that reside within

Table 10.3 Properties of Saturn's largest moons[a]

Name	Mean distance from Saturn (radii)	Orbital period (days)	Mean radius (km)	Mass (10^{19} kg)	Mass density (kg m^{-3})
Mimas	3.08	0.942	198.2	3.75	1150 ± 4
Enceladus	3.95	1.370	252.1	10.79	1608 ± 3
Tethys	4.88	1.888	533.0	61.75	973 ± 4
Dione	6.26	2.737	561.7	109.57	1476 ± 4
Rhea	8.73	4.518	764.3	230.70	1233 ± 5
Titan	20.22	15.945	2575.5	13454.43	1880 ± 4
Hyperion	24.53	21.277	135.0	0.558	542 ± 48
Iapetus	59.01	79.331	735.6	180.58	1083 ± 7
Phoebe	214.97	550.565R	106.6	0.829	1634 ± 46

[a] The orbital distances are given in units of Saturn's equatorial radius, which is 60 268 kilometers and nearly 10 Earth radii. The satellite radii are given in units of kilometers (km). By way of comparison, the mean radius of the Earth's Moon is 1738 kilometers. The mass is given in units of 10^{19} kilograms (kg) – our Moon's mass is 7348 in these units. The mass density is given in units of kilograms per cubic meter (kg m^{-3}).

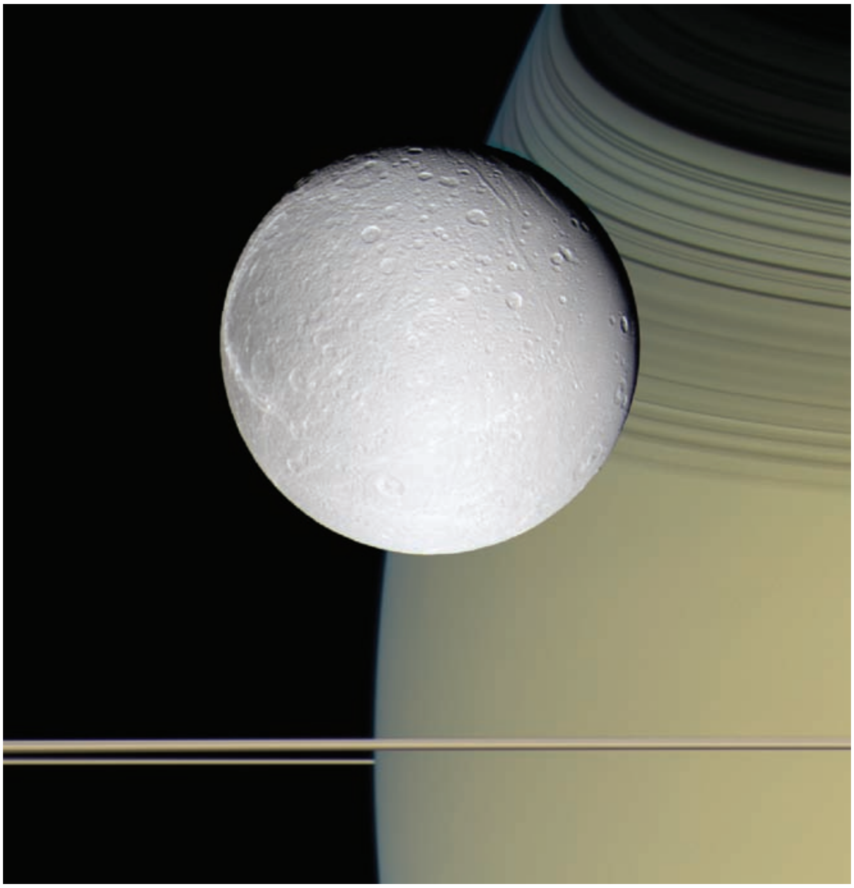


Fig. 10.18 Saturn's moon Dione The surface of pale, icy Dione appears to hover above Saturn's rings seen edge-on (*horizontal stripes near the bottom*), with the tranquil gold and blue hues of the planet in the distant background. Observations at blue, green and infrared wavelengths have been combined to produce this color view, which approximates the scene as it would appear to the human eye. This image was taken on 11 October 2005 from the *Cassini* spacecraft. (Courtesy of NASA/JPL/SSI.)

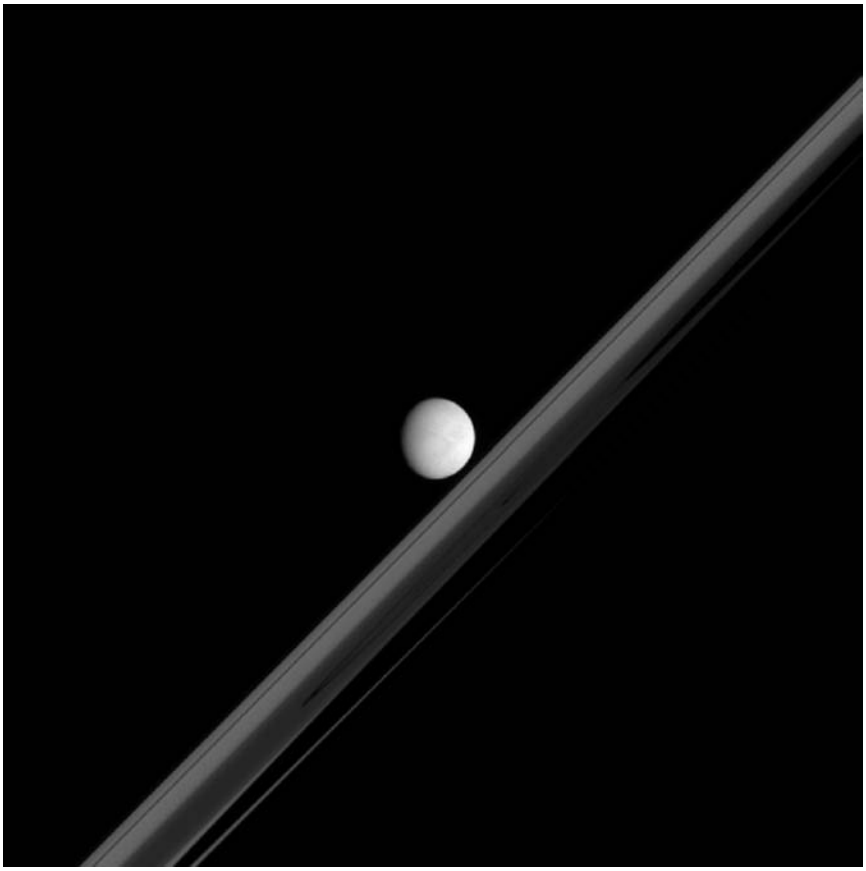


Fig. 10.19 Saturn's bright moon Enceladus Ice-covered Enceladus, about 504 kilometers across, looks suspended in front of rings darkened by Saturn's shadow. From the outer ring edge inward, the observed ring features include the A ring, the Cassini Division and the B ring. The C ring is the dominant, darker region. This image was taken in visible light on 7 March 2005 from the *Cassini* spacecraft. (Courtesy of NASA/JPL/SSI.)

the inner parts of Saturn's satellite system. They are all bright objects; probably composed of ice, and many of them have orbits that are remarkable in one way or another.

At least six of these tiny moons are associated with the rings: Pan, Atlas, Pandora, Prometheus, Janus and Epimetheus. Pan disturbs particles in the A ring to form the Encke Division. Pandora and Prometheus shepherd the F ring; Atlas shepherds the outer margin of the A ring.

Saturn's two co-orbital satellites, Janus and Epimetheus, are even more bizarre. Janus and Epimetheus move in almost identical orbits. The satellite on the inner orbit that is closest to Saturn moves slightly faster, overtaking the outer satellite every four years. But the bodies' diameters are greater than the distance between their orbital paths, so they cannot pass without some fancy pirouetting. They avoid a collision at the last moment by gravitationally exchanging energy and switching orbits. The inner one is pulled by the outer one and raised into the outer orbit, and vice versa. They then move apart, only to repeat this *pas de deux* four years later, and exchange again.

Three so-called Lagrangian satellites move along the orbits of Saturn's larger satellites Tethys and Dione. The satellite Tethys shares its orbit with two small companions, Telesto and Calypso, one about 60 degrees ahead and the other about 60 degrees behind. These two positions, first specified by the Italian-born French mathematician Joseph Louis Lagrange (1736–1813) in the 19th century, are places where the gravitational pull of Saturn and the larger moon are equal. One additional miniature moon, named Helene, shares Dione's orbit, leading it by 60 degrees.

Like Jupiter, the planet Saturn also has a host of small moons discovered from ground-based telescopes.

But there are four fascinating moons of Saturn that have been most extensively studied. They are bright Enceladus, hazy Titan, tumbling Hyperion, and remote, renegade Phoebe, and we now turn to these four, starting with Enceladus.

10.6 Saturn's active water moon Enceladus

Spacecraft view Enceladus close up

Saturn's enigmatic moon Enceladus is small, white and bright, as fresh as new-fallen snow (Fig. 10.19). The diminutive satellite, a mere 504 kilometers across, reflects nearly 100 percent of the incident sunlight, making it one of the brightest known moons at visible wavelengths. Telescopic infrared spectra also indicated a surface of almost pure water ice, which might explain the high reflectivity.

Fig. 10.20 Active Enceladus feeds Saturn's E ring Geysers near the south pole of Saturn's moon Enceladus (*middle of ring*) send water ice and water vapor tens of thousands of kilometers into space, where they are trapped by Saturn's gravity into orbit around the planet, forming the E ring. Saturn's moon Tethys is at the far left, and a background star is seen between Tethys and the ring. The Sun was almost directly behind Saturn when this image was taken, and this backlighting makes the small particles in the E ring appear brighter than they normally are. The dark region extending up from Enceladus follows the moon in its orbit; this hollowed-out core is probably caused by the sweeping action of Enceladus as it moves in the center of the E ring. This image was taken in visible light from the *Cassini* spacecraft on 15 September 2006. (Courtesy of NASA/JPL/SSI.)

When viewed close up from *Voyager 2* in 1981, just a few craters were found on the face of Enceladus, and large, smooth, crater-free regions were discovered, apparently coated with water ice. Parts of the surface contained long cracks and groves, where liquid water might have risen from a warm interior, freezing into smooth ice on the surface. The moon's bright surface suggested that this resurfacing occurred in the fairly recent past, since the ice had not existed long enough to become cratered or darkened.

At the time, it was known that Enceladus orbits Saturn in the middle of the wide, diffuse E ring, discovered in 1966, which is composed of tiny water-ice particles. Theoretical calculations indicated that such a ring will spread out into invisibility over cosmic time intervals, and ought to be replenished in order to survive, so it was hypothesized that Enceladus is feeding the E ring (Fig. 10.20).

Thus, after the *Voyager 2* encounter, scientists postulated that ice, or even liquid water, might be expelled from a currently active Enceladus, accounting for its smooth, highly reflective surface and location near the core of the E ring. But this was an educated guess, an unproven speculation, until the *Cassini* spacecraft passed close enough to test the idea.

Between 2005 and 2010, the *Cassini* spacecraft swooped down as low as 25 kilometers above Enceladus, showing that its south polar landscape exhibits sustained, continuous activity, spewing out ice particles, water vapor and trace amounts of organic chemicals into space. Most of the material does not rise fast enough to escape the moon's gravitational pull, and falls back to cover it with freezing water, but some of the ejected ice has sufficient velocity to escape and go into orbit around Saturn, populating the E ring.

Although the north polar region of Enceladus is pitted with craters, the southern surface is completely without craters and has been twisted and buckled, producing long, parallel, linear troughs, dubbed "tiger stripes", which are stained with dark organic material, run north–south, and are surrounded by roughly circular, circumpolar mountain ridges and valleys (Fig. 10.21).

Active, geyser-like jets arise from the warm tiger-stripe fractures

It is the tiger stripes that mark the location of Enceladus' activity, the places where jets blast, vent and spray out water, lofting it into space somewhat like the Old Faithful Geyser in Yellowstone National Park. These nearly evenly spaced fractures are deep, narrow and long. The V-shaped cracks are about 500 meters deep, with slopes that seem to be coated with ice particles, the fallout of active jets. The fractures are less than 1 kilometer wide, and can extend up to 175 kilometers across the south polar terrain, terminating in hook-shaped bends.

Under the low gravity on Enceladus, the geyser-like jets spew out for thousands of kilometers into space, bending into giant flame-shaped plumes, creating a halo of ice, dust and gas around the moon, and pumping icy material into Saturn's E ring. Water vapor, water ice, powdery snow, carbon dioxide, nitrogen, methane, ammonia, and small amounts of carbon-bearing organic molecules, including acetylene and propane, have been observed in the jets.

Numerous large, small and variable jets have been detected all along the tiger stripes (Fig. 10.22). An individual vent probably stays active until condensing vapor and falling ice particles plug it up and close it off, while the pressure of underground heat forces new vents to open somewhere else along the fracture. The cracks could also open and close under the tidal pull of Saturn's gravity, further controlling the timing of the eruptions. Altogether an endless sequence of jets is produced, varying in space and time with the opening up of new vents and the closing off of old ones.

Infrared detectors aboard the *Cassini* spacecraft have shown that the tiger-stripe sources of the jets glow all along

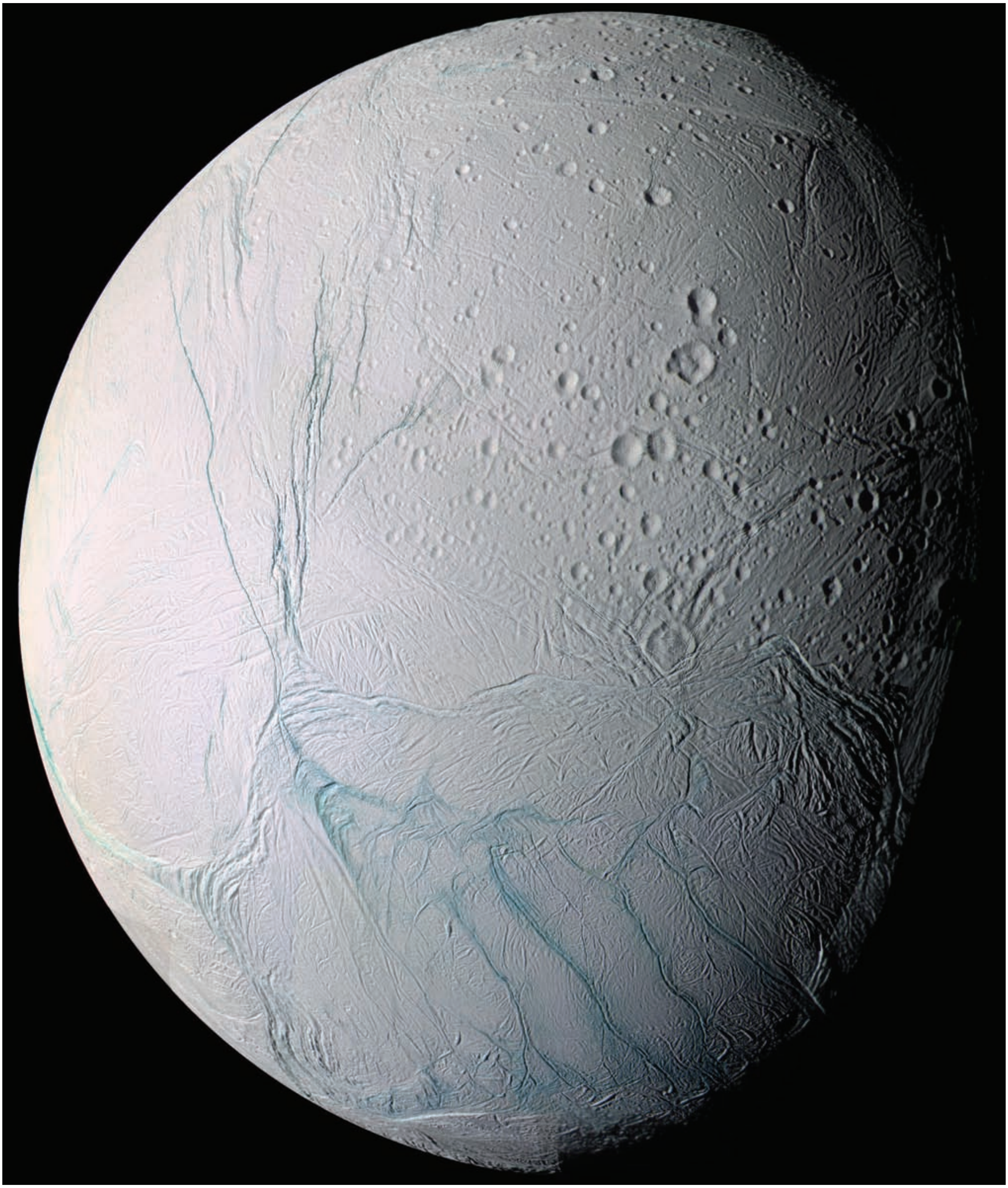


Fig. 10.21 Tiger stripes on Enceladus Saturn's enigmatic moon Enceladus is a jumbled world of fresh snow plains (*middle*), old cratered terrains (*top*), and prominent tiger-stripe fractures (*bottom*, *false color blue*). The fissures spray ice particles, water vapor and organic compounds outward, some of them forming Saturn's E ring and others falling back on the moon. In the mosaic shown here, three prominent tiger stripes extend from the bottom center upwards toward the center. From left to right, they are named Alexandria Sulcus, Cairo Sulcus, and Baghdad Sulcus, the longest tiger stripe. Across the middle of the image, near the northern end of the tiger stripes, 90-degree bends curve along similar paths, starting in a direction parallel to each tiger stripe and then turning perpendicular. Changes in tectonic stresses most likely cause the bends and narrow ridges. This mosaic of images was obtained in visible light from the *Cassini* spacecraft on 21 November 2009. (Courtesy of NASA/JPL/SSI.)

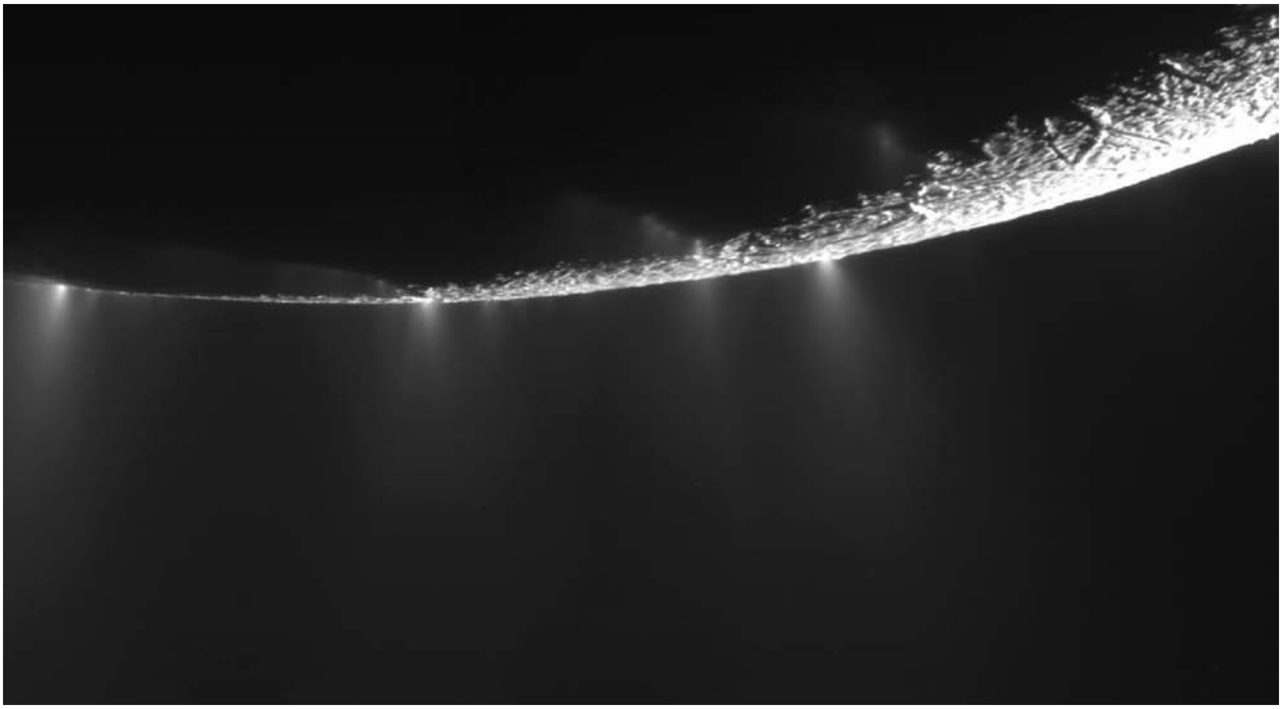


Fig. 10.22 Enceladus vents water jets Dramatic plumes, both large and small, spray out water-ice particles, water vapor and organic compounds from many locations along tiger-stripe fractures near the south pole of Saturn's moon Enceladus. More than 30 individual jets of different sizes can be seen in this image, and more than 20 of them had not been identified before. The south pole of Enceladus lies near the limb in the top left quadrant of the mosaic, near the large jet that is second from left. (Courtesy of NASA/JPL/SSI.)

the fractured lengths. It can remind one of William Blake's (1757–1827) poem, *The Tyger*, with the famous opening line "Tyger! Tyger! Burning bright". Unexpected warmth is found within the giant fractures, with temperatures that can exceed 180 kelvin, more than twice the average temperature of the moon's surface and well above the 70 kelvin that would be expected by heating from sunlight alone. The excess heat appears to be confined within narrow, elongated regions, no more than a kilometer wide and stretching along the fractures.

Why is Enceladus hot inside?

So where does the unexpected warmth in the tiger stripes come from? Enceladus is too small to generate much heat on its own. With a bulk density of 1600 kilograms per cubic meter, it might have a rock core surrounded by a mantle of water ice, but the radioactivity of those rocks cannot produce the observed heat.

There are three possible explanations for the heat, all related to the tidal effects of Saturn's varying gravitational pull on the small moon as it moves along its non-circular orbit. The moon's stretched-out, elliptical trajectory, with an orbital eccentricity of 0.047, is retained by an orbital resonance with the bigger moon Dione, which makes one circuit around Saturn during every two orbits of Enceladus and repeatedly pulls on the smaller satellite.

The varying tidal forces on Enceladus will cause the tiger-stripe fractures to move back and forth, generating heat on their sides as they rub against each other, somewhat like rubbing your hands together to keep warm on a cold winter day. The frictional heat might cause nearby ice to sublimate, or evaporate, into water vapor, which might drag ice particles into space, without the need of any liquid water inside Enceladus.

The tidal flexure of Enceladus might also open and close its fractures over the course of an eccentric orbit, uncovering a buried form of ice called clathrate, which will explosively decompose when exposed to the vacuum of space. Carbon dioxide, methane and nitrogen ought to be released during the explosion, which might propel the water-rich jets outward.

In a third possibility, the tidal squeezing of Enceladus as it moves around its non-circular orbit could make it hot enough inside to melt interior ice. Water vapor produced by this internal heating source could well up in the tiger-stripe fractures, propelling ice particles further out. The temperatures might become high enough at lower depths to produce liquid water inside Enceladus.

The moon is freezing cold on the outside, and under normal circumstances water ice will not melt until it is hotter than 273 kelvin. However, some ammonia has been detected in the jets, and if present in large enough quantities the melting temperature of water might be lowered to about 170 kelvin. Ammonia dissolves in water and acts like the antifreeze in your car, keeping water liquid at a lower temperature than would otherwise be possible.

Although ammonia has not yet been observed in great abundance in the jets, unusually high levels of salt have been detected in ice grains expelled during the eruptions. Scientists conclude that liquid water must be present within Enceladus in order to dissolve enough minerals to account for the levels of salt detected.

A hypothetical sea under the ice-covered south polar region of Enceladus would contain liquid water, organic chemical elements, and an internal source of heat energy, making it a habitable place where living things could survive. It might be analogous to the mid-ocean ridges on Earth's dark sea-floor where exotic creatures thrive on volcanic heat and chemicals in the water, in the complete absence of sunlight. Such a possibility had already been suggested for Jupiter's moon Europa. But of course the possibility of similar things living inside of either Enceladus or Europa is pure unproven, though informed, speculation.

This brings us to Saturn's biggest satellite Titan, 10 times larger than Enceladus. Although liquid water cannot now lap the shores of Titan, it contains seas of liquid methane and ethane.

10.7 Hidden methane lakes and organic dunes on Saturn's moon Titan

Titan's thick, hazy atmosphere

The Dutch astronomer Christiaan Huygens (1629–1695) discovered Titan on 25 March 1655; he was also the first to notice that Saturn has rings. Titan is the largest of Saturn's satellites, the second largest satellite in the solar system, and the only satellite possessing an extensive, dense atmosphere.

The sharp-eyed Catalan astronomer Josep Comas Solà (1868–1937) noticed that Titan's tiny disk is dark at the edges, suggesting that it has an atmosphere. It was confirmed by the Dutch-American astronomer Gerard Kuiper (1905–1973), by spectroscopic observations of methane in Titan's atmosphere.

As the result of space-age investigations, we now know that the surface pressure of Titan's atmosphere is an ear-popping 1.467 bars. That is one and a half times the 1-bar air pressure at sea level on Earth, and equivalent to the pressure experienced by a deep-sea diver at about 6 meters under the ocean's surface.

But it is very cold out there. Besides being about a billion kilometers from the Sun's heat, the surface of Titan lies below a haze which blocks out about 90 percent of the incident light. As a result, the surface temperature is 93.5 kelvin.

When the *Voyager 1* and *2* spacecraft passed behind Titan, as seen from Earth, the homebound radio signal penetrated the atmosphere, permitting an accurate determination of the surface radius from the time the signal disappeared. Under its thick atmosphere, the solid surface of Titan has a mean radius of 2576 kilometers, just slightly smaller than Jupiter's satellite Ganymede, at 2631 kilometers, and larger than the Earth's Moon, with a radius of 1737 kilometers. Titan is also a little larger than the planet Mercury, whose mean radius is 2440 kilometers. Titan is big enough to be a small planet except it is in orbit around Saturn, so Titan is a moon, not a planet.

The trajectories of the *Voyager* and *Cassini* spacecraft have been deflected by a small amount due to Titan's gravitational pull. The size of the deflection permitted an improved determination of Titan's mass. From the mass and radius we can determine the mean mass density, or bulk density, of Titan – 1880 kilograms per cubic meter. That is almost twice the mass density of water ice. If Titan were solid rock, like the Earth's Moon, its average density would be about three times that of water ice. So Titan must be about half rock and half ice.

Precise tracking of *Cassini* during its close encounters with Titan has been used to determine the distribution of materials in the moon's interior, from their gravitational tugs on the spacecraft. The ice and rock are mixed in roughly equal proportions inside Titan, with no separation into distinctive layers other than the outermost 500 kilometers of relatively pure ice.

By way of comparison, Mercury has an average mass density greater than five times that of water. It has a dense iron core in addition to a rocky mantle, and a magnetic field generated by currents in the core. Titan does not have such a core or any detectable magnetic field.

Visible light cannot penetrate the veil of red-orange smoggy haze that covers Titan's surface (Fig. 10.23). In the satellite's dry, cold atmosphere, the smog builds up to an impenetrable haze that extends above the bulk of the atmosphere. On Earth, smog similarly forms by the action of sunlight on hydrocarbon molecules in the air. The urban smog usually forms within a kilometer of the Earth's surface. Titan's atmosphere extends far above its surface because of the high atmospheric pressure and the relatively low mass and gravitational pull of Titan.

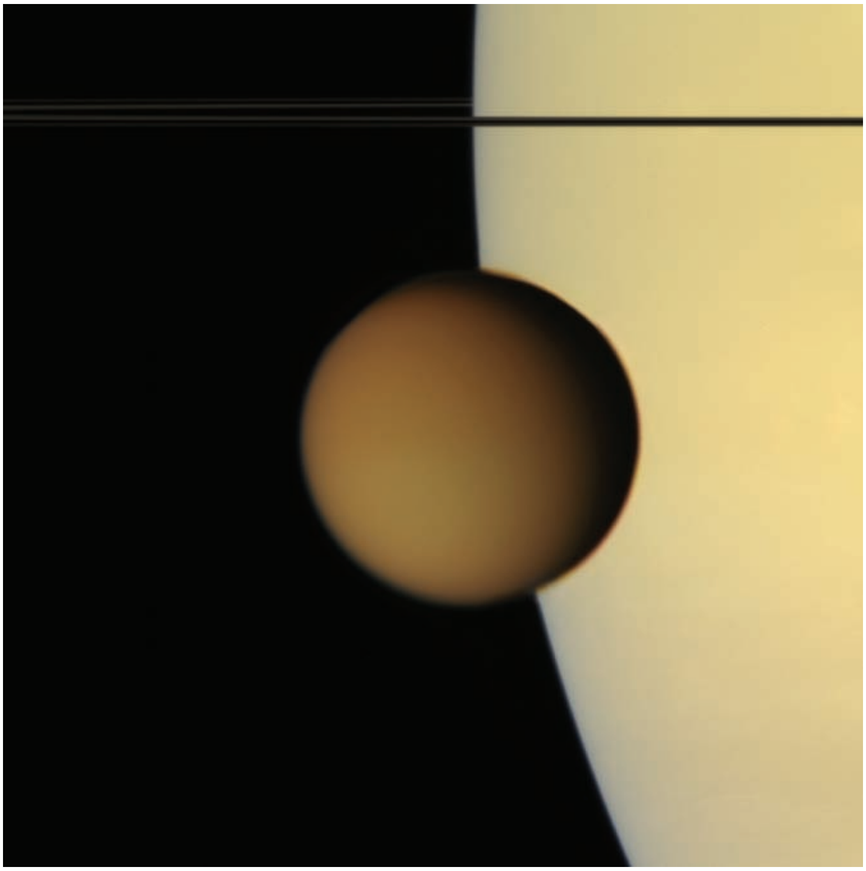


Fig. 10.23 Saturn's hazy moon Titan
The murky orange disk of Titan passes in front of Saturn and just below the planet's rings. Titan's photochemical smog completely obscures the moon's surface. High-altitude hazes are visible against the disk of Saturn, attenuating light reflected by the planet. Images taken at red, green and blue wavelengths have been combined to create this natural-color view, taken on 1 August 2007 from the *Cassini* spacecraft. (Courtesy of NASA/JPL/SSI.)

Titan's atmosphere extends at least 10 times further than does our own. Moreover, the high-flying stratospheric haze material, named tholin for the Greek word for "muddy", lies at altitudes above the surface of greater than 1000 kilometers. The reddish tholins are very large, complex organic molecules thought, by some, to be the chemical precursors to life.

Instruments aboard the *Voyager 1* spacecraft revealed several fascinating things about Titan, though its surface remained largely hidden. They found that the dominant constituent of the thick, heavy atmosphere is molecular nitrogen, now pegged at 98.4 percent of the atmosphere, with the remaining 1.6 percent composed of methane and trace amounts of hydrocarbons.

So, nitrogen molecules account for the bulk of Titan's atmosphere as they do on Earth – 77 percent of our air is molecular nitrogen. But, unlike Earth, the atmosphere of Titan contains no molecular oxygen, which accounts for 21 percent of the Earth's atmosphere. The freezing temperature on Titan is way too low for any living things, such as plants, to supply oxygen. Titan therefore has a hydrogen-rich atmosphere, rather than an oxygen-rich one.

High above Titan's surface, abundant nitrogen and methane molecules are being broken apart continuously by the Sun's energetic ultraviolet light and by the bombardment of electrons trapped in Saturn's magnetic environment. Some of the fragments then recombine to form more complex molecules that have been detected in small amounts by the *Voyager 1* infrared spectrometers (Fig. 10.24). Almost 20 organic gases were identified in Titan's atmosphere. In addition to methane (CH_4), the list includes hydrocarbons like ethane (C_2H_6), acetylene (C_2H_2), and propane (C_3H_8), and nitrogen compounds such as hydrogen cyanide (HCN). Many of these molecules can join together in chainlike polymer structures that contribute to Titan's dark, smoggy haze.

Although it could not see beneath the smog surrounding Titan, the radio signals sent home from *Voyager 1* have been used to infer the vertical structure of the pressure and temperature down to the surface (Fig. 10.25). Titan's temperature profile is very similar to that of the Earth's atmosphere, where the temperature initially drops with increasing altitude above the surface and then rises again, so Titan's atmosphere has the equivalent of our ground hugging troposphere and higher stratosphere.

On Earth, it is ozone that absorbs ultraviolet sunlight, making the stratosphere. On Titan, it is a photochemical haze, produced by sunlight's destruction of methane. Also unlike Earth, the temperatures in Titan's atmosphere are everywhere below the freezing point of water, and peak at 175 kelvin at about 40 kilometers above the surface.

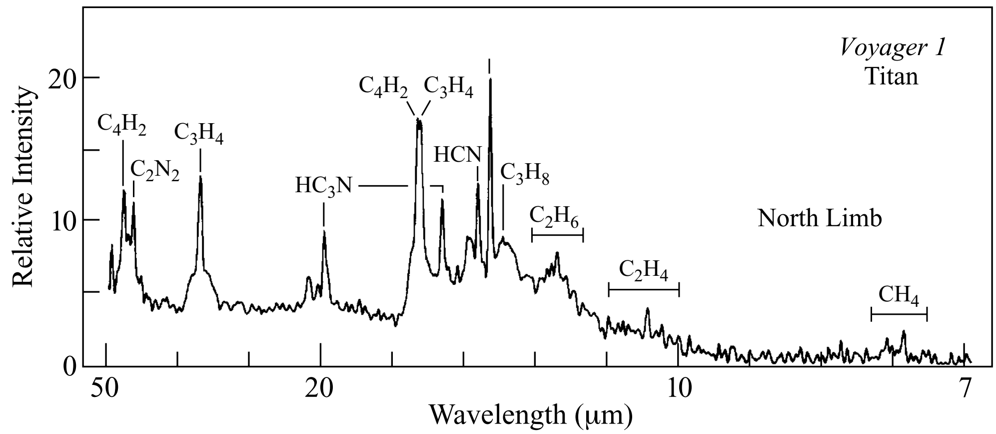


Fig. 10.24 Molecules in Titan's atmosphere Emission features in the infrared spectrum of Titan's reflected sunlight identify the molecular constituents of its atmosphere. Sharp peaks in this spectrum, acquired from the *Voyager 1* spacecraft in 1980, are attributed to methane (CH_4), acetylene (C_2H_2), ethane (C_2H_6), and more complex hydrocarbon molecules, as well as nitrogen compounds. The wavelength is given in units of microns (μm), or 10^{-6} meters.

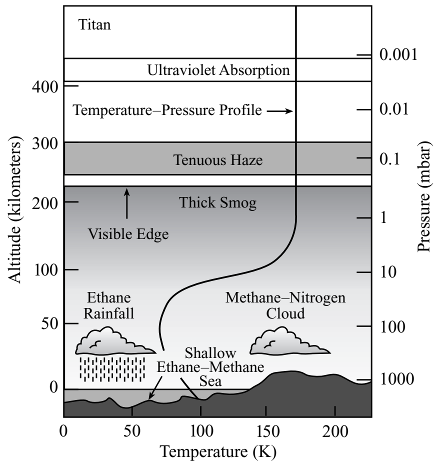


Fig. 10.25 Model of Titan's atmosphere A study of the bending and fading of homebound radio signals when the *Voyager 1* spacecraft passed behind Titan led to this plot of the temperature and pressure in the satellite's atmosphere. The temperature (*bottom axis*) decreases with height until about 40 kilometers altitude, and then increases rapidly at higher altitudes (*left axis*). The entire atmosphere is well below the freezing temperature of water, at 273 kelvin, but the lower atmosphere is just warm enough to allow the condensation of liquid nitrogen. The high-altitude smog on Titan covers clouds of methane, and liquid ethane and methane probably rain down to the surface. The pressure (*right axis*) is given in units of millibars or 0.001 bars, and it reaches 1500 millibars or 1.5 bar near the surface of Titan. The air pressure at sea level on Earth is 1.0 bar. The pressure axis is logarithmic, which means that the numbers on the altitude axis are not uniformly spaced.

Focus 10.2 Titan could be an early Earth in a deep freeze

The chemistry in Titan's hydrogen-rich atmosphere may be similar to that in Earth's atmosphere several billion years ago, before living things released molecular oxygen into the air and modified it. So Titan could serve as a time machine, taking us back to a simpler era on Earth before life began to contaminate the planet. Titan could even provide clues to how life got started when the Earth was young. Nevertheless, the exceedingly low surface temperature rules out current life on Titan.

All the life that we know about depends on molecules that contain carbon and hydrogen atoms, and such hydrocarbon molecules have been found in Titan's atmosphere. The chemical study of these compounds is known as organic chemistry – but it has nothing to due with organic foods grown without artificial fertilizers. And on Titan the organic chemistry is going on without concurrent life.

So Titan is a frozen moon that resembles the early Earth in a deep freeze. About seven billion years from now, the Sun will near the end of its life and swell up to become a bright giant star. The intense heat from the aged and swollen Sun will warm Titan's surface and may bring it to life. The moon will become an oasis of liquid water and organic chemicals, ready to initiate life.

The surface temperature on Titan, of 93.5 kelvin, is far too cold to permit life, but its atmosphere may nurture chemical reactions similar to those at work on Earth before life began there (Focus 10.2). But Titan's temperature is close to the triple point of methane, which would permit it to exist in solid, liquid and gaseous form, just as Earth's temperature is near the triple point of water. The methane can condense in Titan's cold atmosphere to produce thick clouds that lie beneath the haze. Infrared observations that penetrate the smog suggest the presence of short-lived methane clouds in Titan's lower atmosphere, which form briefly and irregularly. Since the atmosphere is not fully saturated with methane, there cannot be extensive oceans of pure methane on the surface, but both ethane and methane can rain out of the atmosphere. They can exist as a liquid rather than a solid at the surface temperature of 93.5 kelvin. Evaporation of the liquid seas can re-supply the hydrocarbons to the atmosphere, completing the cycle.

Thus, we might expect seas, lakes and ponds of liquid hydrocarbons on Titan, consisting of ethane and methane, and the *Cassini* spacecraft and its companion *Huygens Probe* have now confirmed their existence.

Huygens Probe lands on Titan

On 14 January 2005, the NASA-built *Cassini* spacecraft dropped the European-built *Huygens Probe* on a 2.5-hour parachute descent through Titan's atmosphere to its long-hidden surface. This was the first time that a spacecraft had touched down on another planet's moon. The probe was designed to survive the impact and splash down on a liquid surface, sending back data for no more than three hours, but no ocean was in sight.

Instruments on the descending probe obtained images of dark, snaking and branching features that merge into a large channel, suggesting methane flows that may have once been fed by methane rainfall. Methane rain and drizzle were, in fact, detected on the way down, but the probe landed in a damp riverbed at equatorial latitudes, rather than a lake or sea. It found only small rocks and pebbles, perhaps rounded by tumbling methane rivers in the past.

The *Huygens Probe* could see only a small section of Titan, but radar images from the *Cassini* spacecraft, still in orbit around Saturn, revealed long, deep channels extending over great distances in other locations, meandering along with few tributaries like some lazy terrestrial rivers of water.

Scientists had perhaps been misled by images taken from afar, which showed dark regions straddling Titan's equator. They were once thought to be possible methane seas, but they contain no more liquid than water on the maria of Earth's Moon. Nevertheless, the long-anticipated liquid lakes were spotted in the north polar regions of Titan.

Lakes of liquid methane and ethane

Radar observations from the *Cassini* spacecraft in 2005 to 2010 have revealed numerous flat, dark features on Titan, which were first reported, with coarser resolution, in 2003 by radio astronomers using the ground-based Arecibo telescope. The radio waves can penetrate the haze and smog to obtain images of the moon's surface. Dark radar images correspond to weaker radar echoes and a smoother terrain, while bright regions correspond to stronger echoes and a rougher surface.

Radar images of parts of the northern regions of Titan are completely black, reflecting essentially no radar signal, and are therefore extremely smooth. They have been attributed to lakes of liquid methane, possibly with ethane dissolved in them. The numerous large lakes discovered by *Cassini* radar are found mainly at northern latitudes, which was then in winter and 2 to 3 kelvin colder than the equator. Since ethane and methane are both highly flammable, much of the satellite could go up in flames, but it won't ignite because of the lack of molecular oxygen in the atmosphere.

Some of the lakes are many tens of kilometers in length, and their perimeters are reminiscent of the shores of terrestrial lakes (Fig. 10.26). Clouds that rain methane have been observed above some of the lakes, plausibly filling them in season. These liquids are also thought to carve the meandering rivers and channels spotted elsewhere on the moon's surface.

Though less common, the lakes also form in Titan's south polar regions. Low-altitude passes over these regions by *Cassini* have confirmed the presence of liquid ethane in one of the largest southern lakes, Ontario Lactus, by the way it absorbs and reflects infrared radiation. Observations of this lake's sloping, receding shoreline suggest that the lake is evaporating in the summer warmth. Seasonal variations of the precipitation and evaporation of methane and ethane, similar to those of water lakes on Earth, might explain why the lakes in Titan's northern high latitudes were 20 times more common than lakes in the southern high latitudes, with rainfall filling the northern lakes in winter.

Organic dunes

Titan's entire surface is not wet. At low latitudes near the moon's equator, the *Cassini* radar instrument detected vast dune fields instead of lakes. The giant, rippled dunes are aligned in the longitudinal direction, suggesting that they are shaped by strong winds blowing mainly from east to west. The linear dark streaks extend hundreds of kilometers and rise to a few hundred meters, with strips between the dunes that are nearly free of dark material. They bear a striking resemblance to sand dunes on Earth, but instead of sand Titan's dunes are made out of solid organic grains. The tiny ice particles are coated with organic substances, and likely derive from organic chemicals in Titan's smoggy skies.

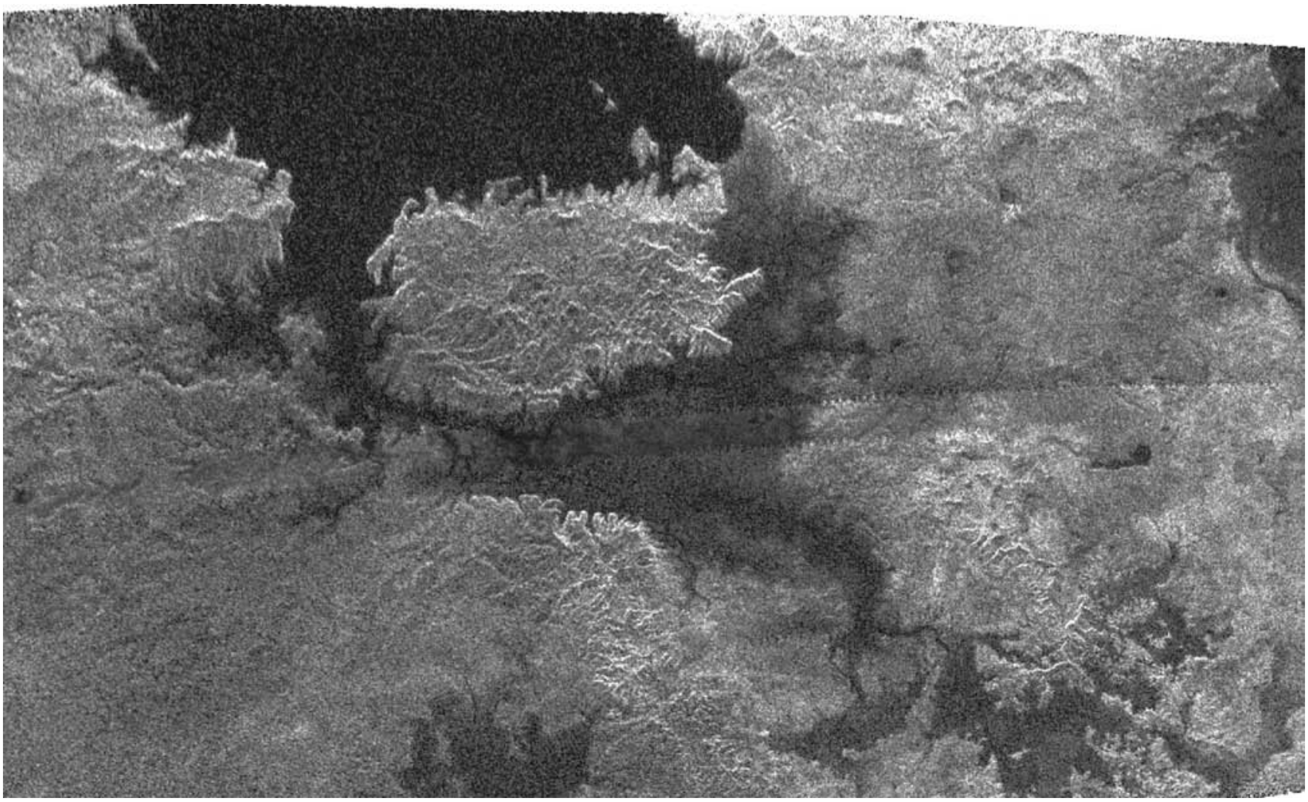


Fig. 10.26 Lakes of methane and ethane on Saturn's moon Titan A radar image shows a bright, central island in a dark, smooth lake, surrounded by a bright shoreline with numerous inlets. The island is about 90 kilometers across. This is one of many large lakes formed at high latitudes on Titan, with more in the northern polar regions than the southern ones. Seasonal rains of liquid methane and ethane probably fill these lakes. The radar instrument aboard *Cassini* obtained this image during a close south polar flyby on 22 February 2007. (Courtesy of NASA/JPL.)

10.8 Alien worlds, distant ring

The *Cassini* spacecraft has also zeroed in on one of Saturn's most unusual, mid-sized moons, named after the god Hyperion. Attention was drawn to this alien world in the 1980s after *Voyager 2* images indicated that it is one of the most irregularly shaped, non-spherical bodies in the solar system, and it was realized that Hyperion's rotational period is not constant. The strange moon also has a very low mean mass density of about 540 kilograms per cubic meter, indicating that it is probably composed of water ice and filled with empty spaces that occupy about half its volume.

Many natural satellites are locked into synchronous rotation through tidal interaction with their planet, including the Earth's moon, the four Galilean satellites, and most of the large or mid-sized moons of Saturn. The rotation period of these satellites is equal to their orbital period, so the same side of the moon always faces their planet. The exception is Hyperion, whose rotation is chaotic. As shown by Stanton J. Peale (1937–) and Jack Wisdom (1953–) in the mid-1980s, strong gravitational torques on the asymmetric satellite, coupled with its large orbital eccentricity and resonance with Titan, cause Hyperion to tumble in a random manner. Its axis of rotation wobbles so much that its orientation in space is unpredictable.

Even more fantastic aspects of Hyperion were revealed when *Cassini* was targeted to pass within 500 kilometers of the moon, on 26 September 2005. A combination of images taken at different wavelengths reveals a surface that is covered with deep, sharp-edged craters with dark material at their bottoms (Fig. 10.27). The largest crater on Hyperion is about 122 kilometers across, or nearly half the moon's average diameter of 270 kilometers, perhaps lending support to the idea that Hyperion is an impact fragment of a former larger body.

Phoebe is another irregular satellite, for it moves around Saturn in the backward retrograde direction and is inclined by 152 degrees to Saturn's equator. For more than 100 years Phoebe was Saturn's outermost known moon, about four times more distant from the planet than Iapetus, the next closest moon, which orbits very nearly in the plane of Saturn's equator.

Fig. 10.27 Saturn's odd moon Hyperion This stunning view of Saturn's asymmetric, tumbling, impact-cratered moon Hyperion reveals numerous sharp-edged craters, which make the moon look like a giant sponge. Hyperion's unusual appearance can be attributed to the fact that it has an exceptionally low mean mass density, of about 540 kilograms per cubic meter, with a large porosity of 0.46 and a weak surface gravity. In other words, the moon seems to contain a lot of empty holes, somewhat like a sponge. This processed, false-color view combines images taken at infrared, green and ultraviolet wavelengths from the *Cassini* spacecraft during a close flyby on 26 September 2005. (Courtesy of NASA/JPL/SS.)

Fig. 10.28 Saturn's retrograde moon Phoebe This image of Phoebe suggests that the moon may be an ice-rich body coated with a thin layer of dark material, resembling the nucleus of comets. Small bright craters in the image are probably young features. When impacting projectiles slammed into the surface of the moon, the collisions excavated fresh, bright material, probably water ice, underlying the surface area. Dark material on some crater walls appears to have slid downwards, exposing more light-colored material. Phoebe orbits Saturn in the backward, retrograde direction to all of the other mid-sized satellites of the planet, and the moon's dark and irregular cratered surface, retrograde orbit, and low mean mass density suggest that Phoebe was once part of the Kuiper belt of icy comets beyond Neptune before passing near Saturn and being captured by its gravity. This mosaic of two images was acquired from the *Cassini* spacecraft during its Phoebe flyby on 11 June 2004. (Courtesy of NASA/JPL/SSI.)

When approaching Saturn in 2004, the *Cassini–Huygens* spacecraft first encountered outlying Phoebe, providing an opportunity to test the mission's imaging equipment before traveling on to encounter the planet. The high-resolution images reveal Phoebe to be a heavily cratered body with large variations in brightness, suggesting that bright water ice lies beneath a thin blanket of dark surface deposits (Figs. 10.28, 10.29). Although once believed to be a captured asteroid, scientists now speculate that Phoebe is a former Centaur, one of a number of icy objects originally residing in the distant Kuiper belt and now orbiting the Sun between Jupiter and Neptune.

Phoebe circles Saturn within an enormous ring in the far reaches of the planet's realm and way outside its Roche limit. The infrared emission of this ring was discovered in 2009 using the *Spitzer Space Telescope* (Fig. 10.30). The ring is tilted 27 degrees from Saturn's equatorial plane, and all the other rings, extends from 128 to 207 times the radius of Saturn, and is about 20 times as wide as the diameter of the planet. The diffuse ring is thought to originate from micrometeorite impacts of the surface of Phoebe.

Fig. 10.29 Bright and dark layers on Phoebe Saturn's mid-sized, retrograde moon, Phoebe, is most likely an ice-rich body overlain with a thin layer of dark material. The sharply defined crater at *above center* exhibits two or more layers of alternating bright and dark material. The layering might have occurred during the crater formation, when material ejected from the crater buried the pre-existing surface that was itself covered by a relatively thin, dark deposit over an icy mantle. This image was taken from the *Cassini* spacecraft during its Phoebe flyby on 11 June 2004. (Courtesy of NASA/JPL/SSI.)

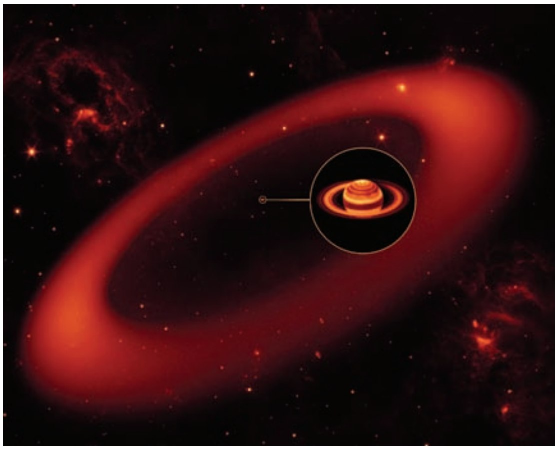


Fig. 10.30 Saturn's enormous infrared ring This artist's conception illustrates the infrared glow of cold dust particles in Saturn's largest ring, discovered using the *Spitzer Space Telescope* in 2009. The very tenuous collection of ice and dust particles is spread out in an enormous belt at the far reaches of Saturn's system, with an orbit tilted 27 degrees from the main ring plane. The bulk of its material starts about 6 million kilometers away from the planet, and extends outward another 12 million kilometers. The ring's diameter is equivalent to roughly 300 times the diameter of Saturn. The planet appears as just a small dot in the middle of this portrayal. The inset shows an enlarged image of Saturn, as seen by the W. M. Keck Observatory at Mauna Kea, Hawaii, in infrared light. Saturn's retrograde moon Phoebe circles within the newfound ring, and is likely to be the source of its material. Dark material dislodged from Phoebe may also be the source of the dark side of Saturn's next innermost moon, two-faced Iapetus. (Courtesy of NASA/JPL-Caltech/Keck.)

The new ring lends support to the idea there is a connection between Phoebe and the dark side of its neighboring moon Iapetus. *Voyager 2* images indicated that two-faced Iapetus is a divided world, with a heavily cratered side as bright as ice and another side as dark as asphalt or coal. The side of Iapetus that faces forward in its orbit is apparently being darkened in some mysterious way, and the darkening particles may be coming from Phoebe's ring. Once deposited, the dark material absorbs sunlight and warms the surface of Iapetus, perhaps vaporizing and releasing water ice that moves away from the dark side.

This concludes our survey of Saturn, the most distant planet known to the ancients. We will now travel out beyond this enchanting world to the next wanderers, Uranus and Neptune.

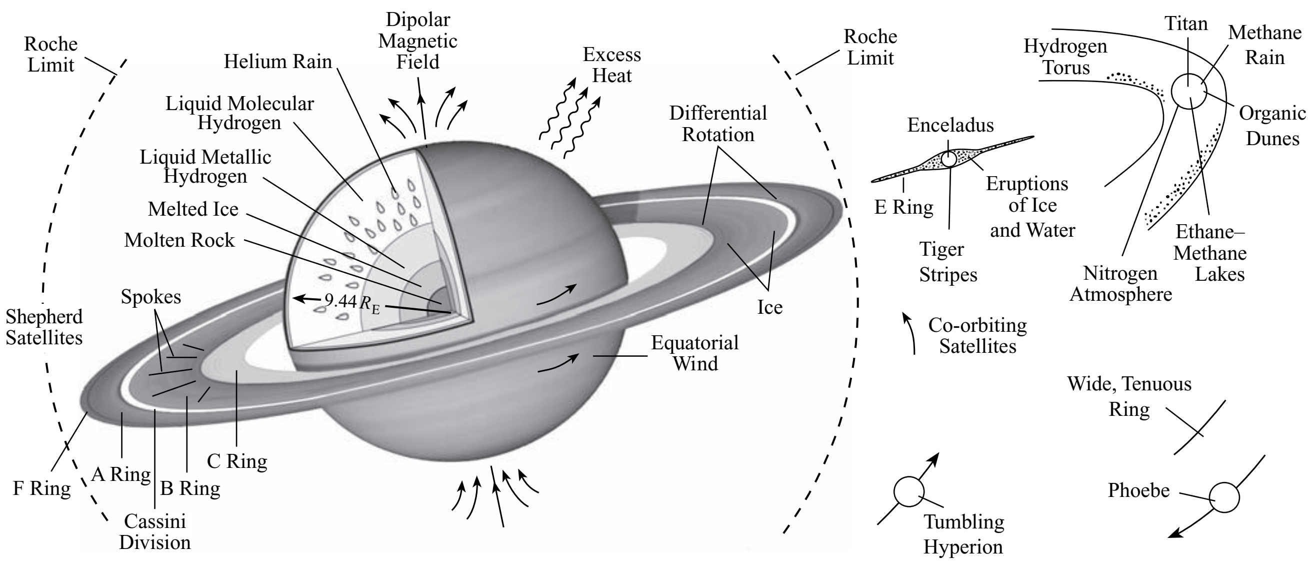


Fig. 10.31 Summary diagram

11 Uranus and Neptune

- Uranus and Neptune were unknown to ancient astronomers, and were not discovered until after the invention of the telescope.
- Uranus is just barely visible to the unaided eye, and Neptune requires a telescope to be seen.
- The blue-green, turquoise color of Uranus and the indigo blue color of Neptune come from methane in their clouds.
- Uranus and Neptune have a similar size, mass, and bulk composition, and they are both much smaller, less massive, and denser than the other two giant planets, Jupiter and Saturn.
- In contrast to all the other planets in the solar system, Uranus is tipped sideways so its rotation axis lies nearly within the planet's orbital plane, leading to extreme seasonal variations in solar heating of the planet's polar regions.
- Although Uranus apparently has no strong internal source of heat, Neptune radiates 2.7 times the energy it absorbs from the Sun. The source of Neptune's excess energy is most likely heat left over from the planet's formation.
- The cloud bands and winds on Uranus blow parallel to the planet's equator, apparently controlled by the planet's rapid spin rather than by direct heating from the Sun.
- Despite receiving a relatively small amount of sunlight compared to the other major planets, Neptune's atmosphere is surprisingly active and dynamic, with large storm systems and high-speed winds that may be driven by internal heat.
- As the southern hemisphere of Neptune turned slowly toward the Sun's faint heat, the temperature of the planet's south polar region increased by just 10 kelvin and methane storm clouds became more frequent in the southern hemisphere.
- Unlike Jupiter and Saturn, there is no liquid metallic hydrogen inside Uranus and Neptune, but they both contain deep atmospheres of molecular hydrogen.
- The internal structure of Uranus and Neptune includes vast internal oceans of water, methane and ammonia "ices", melted at the high temperatures inside.

- The magnetic fields of Uranus and Neptune are askew, tilted from their rotation axes; rotation-driven currents in internal shells of ionized water could generate these magnetic fields.
- The austere, skeletal rings of Uranus are very narrow and widely spaced from each other, and made of very dark material.
- Shepherd satellites are most likely responsible for the narrowness of the rings of Uranus and Neptune.
- The rings around Uranus are not quite circular, do not lie exactly in Uranus' equatorial plane, and vary in width; these irregularities are attributed to the gravitational interaction of ring particles with small nearby moons.
- The material in one narrow ring around Neptune has been concentrated into three clumps, probably by the gravity of a nearby moon or moons.
- The sparse rings around Neptune contain no more material than that found in a single small moon only a kilometer across.
- Most of Neptune's rings that we see now will probably be ground into dust by collisions and meteoritic bombardment in a few hundred million years, eventually being consumed by their central planet and vanishing from sight. But the rings can easily be replaced by debris blasted off small moons already embedded in them.
- The five major moons of Uranus are dark and dense, made up predominantly of rock and water ice.
- Miranda, the innermost mid-sized satellite of Uranus, exhibits a bizarre variety of surface features that suggest repeated violent impacts in the past. It may have been shattered by catastrophic collisions and reassembled, or it became frozen in an embryonic stage of differentiation.
- Uranus' large moons all revolve in the planet's equatorial plane, almost perpendicular to the planet's orbital plane, in circular synchronous satellite orbits with the same side always facing Uranus.
- Neptune's largest satellite, Triton, revolves about the planet in a direction opposite to that in which Neptune rotates.
- Triton has a very tenuous, nitrogen-rich atmosphere, bright polar caps of nitrogen and methane ice, frozen lakes of water flooded by past volcanoes of water ice, and tall geysers that may now be erupting on its surface.
- Triton may have formed in orbit around the Sun and was subsequently captured by Neptune, whose tidal forces kept Triton molten for much of its early history. These tides are now pulling the satellite toward a future collision with the planet.

11.1 Fundamentals

Planetary twins

Saturn was the most distant planet known to the ancients. Uranus and Neptune are both so far away, and so faintly illuminated by the Sun, that telescopes were required to discover them. One can just barely discern Uranus with the unaided eye, but few astronomers have seen it without a telescope. Neptune cannot be seen without the aid of a telescope. William Herschel (1738–1822) discovered Uranus in 1781 during his telescopic survey of the faint stars located near bright ones, and Neptune was found in 1846, as the result of a mathematical

Table 11.1 Some comparisons of Uranus and Neptune[a]

	Uranus	Neptune
Mass (Earth mass)	14.53	17.14
Equatorial radius (Earth radius)	3.98	3.91
Bulk density (kilograms per cubic meter)	1270	1638
Sidereal rotation period (hours)	17.24	16.11
Sidereal orbital period (Earth years)	84	165
Mean distance from Sun (AU)	19.19	30.06
Outer atmosphere	82.5 percent hydrogen	80.0 percent hydrogen
	15.2 percent helium	18.5 percent helium
	2.3 percent methane	1.5 percent methane
Energy balance	Less than 1.4	2.7 ± 0.3
Effective temperature	59.3 kelvin	59.3 kelvin
Temperature at one-bar level	76 kelvin	73 kelvin
Central temperature	5000 kelvin	5000 kelvin
Magnetic dipole moment[b]	50 D_E	25 D_E
Equatorial magnetic field strength	0.23×10^{-4} tesla	0.14×10^{-4} tesla

[a] The Earth's mass is 5.9722×10^{24} kilograms, the Earth's equatorial radius is 6378 kilometers, the astronomical unit (AU) is the mean distance between the Earth and the Sun with a value of 1.496×10^{11} meters. The energy balance is the ratio of total radiated energy to the total energy absorbed from sunlight, the effective temperature is the temperature of a black body that would radiate the same amount of energy per unit area, and a pressure of one bar is equal to the atmospheric pressure at sea level on Earth.

[b] The magnetic dipole moment is the product of the equatorial magnetic field strength and the cube of the planet's radius, given here in units of the Earth's magnetic dipole moment $D_E = 7.91 \times 10^{15}$ tesla meters cubed.

prediction based on its gravitational effect on the motion of Uranus.

These two outermost major planets are considered giant planets, but they are much smaller, less massive and denser than the other two giants, Jupiter and Saturn. Uranus and Neptune are just four times bigger than the Earth, only about one-third the size of Jupiter, and less than half the size of Saturn. The mass of Uranus and Neptune has been inferred from the orbital periods and distances of their largest moons. They have just 14.53 and 17.14 times the Earth's mass, for Uranus and Neptune respectively, which is about the same mass as the ice–rock cores of Jupiter or Saturn.

Uranus is about 19 times as far away from the Sun as the Earth is, and Neptune is about 30 times as distant. As a result, it takes 84.0 Earth years for Uranus to complete one revolution about the Sun and nearly twice that for Neptune.

These two distant planets remain little more than dim, fuzzy spots of light in even the most powerful telescope. From the Earth, the planet Uranus subtends an angle of just 3.5 seconds of arc, and Neptune 2.0. Since the Earth's atmosphere blurs features smaller than about 1.0 seconds of arc, ground-based observers cannot distinguish features in the outer atmospheres of Uranus and Neptune.

One can still infer enough about Uranus and Neptune from telescopic observations to know that they have similar bulk physical properties. The size, mass, composition and rotation of Uranus and Neptune are in fact so similar that they are often called planetary twins (Table 11.1). From each planet's mass and size, we calculate its bulk density, which is intermediate between the low-density giants, Jupiter and Saturn, and the dense rocky Earth or Mars. Most of Uranus and Neptune must therefore be composed of something less dense than rock, but more substantial than the hydrogen and helium that dominate the composition of Jupiter and Saturn. The main ingredients of Uranus and Neptune are probably liquid water and other melted ices.

The blue-green, turquoise color of Uranus and the deeper indigo blue of Neptune are attributed to methane. At the low temperatures prevailing at their cloud tops, the methane condenses to form a top layer of clouds made of methane ice crystals. And since methane absorbs red light quite strongly, the sunlight reflected off the thick, deep clouds of Uranus and Neptune has a blue color.

The blue cloud deck of methane ice forms where the atmospheric pressure is about one bar, or roughly the same as the air pressure at sea level on Earth. At this level, the equatorial radius of Uranus is 25 559 kilometers while that of Neptune is 24 764 kilometers. Rapid internal rotation produces an equatorial bulge on both planets; the polar radius is 586 kilometers shorter on Uranus and 424 kilometers on Neptune.

The sidereal rotation periods of Uranus and Neptune lie between the roughly 10-hour day of Jupiter and Saturn and the 24-hour rotation period of Earth and Mars. Periodic variations in the radio emission of Uranus and Neptune, detected from the *Voyager 2* spacecraft, indicate that their magnetic fields, which are anchored deep inside the planets, rotate once with respect to the stars every 17.24 hours and 16.11 hours, respectively.

Thus, Uranus and Neptune bear an uncanny resemblance to each other, but there are two remarkable differences. Uranus is tipped sideways and has no significant heat inside, while Neptune is more upright and has a lot of internal energy.

Uranus is tipped on its side and has no strong source of internal heat

Blue-green Uranus is tipped sideways, with its poles where its equator should be. Unlike all the other planets, whose rotation axes are roughly perpendicular to their orbital plane, the ecliptic, the rotation axis and poles of Uranus lie almost within the ecliptic. The equatorial plane of Uranus is inclined 97.9 degrees from its orbital plane, with a tilt that is just a bit more than a right angle.

This knowledge comes not from watching Uranus rotate, but instead from observing the orbits of its major moons. The orbits are all circular, and they lie in one plane, which is turned at right angles to the plane of Uranus' orbital motion about the Sun. As a result, the moons form a bullseye pattern, revolving around Uranus like a Ferris wheel. Since these satellites should be orbiting within the plane of Uranus' equator, the entire planet has to be tipped on its side (Fig. 11.1). One speculation is that Uranus was knocked sideways during a glancing collision, between the planet and another massive body, perhaps when the planet was still forming.

Because the rotational axis of Uranus lies near its orbital plane, first one pole and then the other points toward the Sun as the planet slowly progresses around its orbit. Each pole faces the Sun for about 21 Earth years, with a corresponding period of darkness at the opposite pole. When either pole points at the Sun, an imaginary observer near that pole would never see the Sun set. Terrestrial astronomers see the moons of Uranus moving in circles around the planet. Between the long summer and winter at each pole, the equator turns toward the Sun, and we observe the moons traveling vertically up and down as they move around the equator.

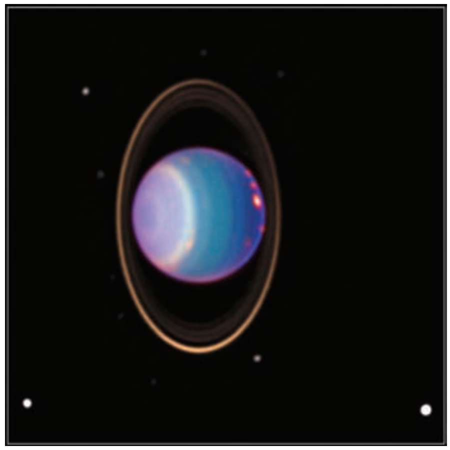


Fig. 11.1 The tilted planet Thin, spidery rings and several small moons encircle Uranus. The planet is tipped on its side, so its equator, rings and direction of rotation are almost perpendicular to the plane of the planet's orbital motion around the Sun. Storm clouds, found in the northern upper atmosphere (*right*) circle the planet at more than 500 kilometers per hour. The clouds are tinted pink in this false-color image to show that it was taken at infrared wavelengths, but the clouds are white in visible light. This image was taken on 8 August 1998 from the *Hubble Space Telescope*. (Courtesy of NASA/JPL/STScI.)

When *Voyager 2* arrived at Uranus, on 24 January 1986, the spacecraft's infrared detectors found that the planet is radiating about as much energy as it receives from the Sun. This means that Uranus lacks a strong internal heat source, in contrast to Jupiter and Saturn that produce heat in their centers. These two giants each radiate away about twice as much energy as they absorb from the Sun. But like Jupiter and Saturn, and unlike Uranus, the planet Neptune glows in the infrared with its own internal heat discovered when *Voyager 2* flew past Neptune in August 1989, twelve years after launch.

The temperature at Neptune's cloud tops is 59.3 kelvin, about the same as that of Uranus. But since Neptune is 50 percent farther from the Sun, it should have been a lot colder than Uranus; the temperature that would result from sunlight alone is 46 kelvin at the cloud tops of Neptune. The hotter measured temperature implies that Neptune radiates 2.7 times as much energy as it absorbs from the Sun, and its outer atmosphere must therefore receive energy from the interior.

11.2 Storm clouds on the outer giants

The outer atmospheres of Uranus and Neptune are quite similar in composition to those of Jupiter and Saturn.

Molecular hydrogen accounts for 82.5 percent of the outer atmosphere of Uranus, followed by 15.2 percent helium atoms and 2.3 percent methane. Neptune has roughly comparable amounts of these elements in its outer atmosphere. Although they are mostly composed of the light gases, hydrogen and helium, it is the methane in the colder cloud levels that accounts for their colors.

Mild weather on Uranus

At the low temperatures prevailing near the top of Uranus' atmosphere, methane gas freezes and forms a methane cloud deck; haze particles are also formed there due to the action of ultraviolet sunlight on methane. The haze and methane clouds hide the lower atmosphere from view. Ammonia and water clouds probably form deeper in the atmosphere and are difficult to see. On warmer Jupiter and Saturn the topmost light-colored clouds that we see are composed of ammonia ice crystals rather than frozen methane.

With no appreciable heat rising from the interior to drive the weather system, Uranus seemed to present a dull and placid face to the world. Looking at Uranus was something like gazing down into a bottomless ocean. Yet some zonal banding was extracted from the *Voyager 2* images and subsequently confirmed using the Earth-orbiting *Hubble Space Telescope* (Fig. 11.2) and the ground-based Keck II telescope. So storms and clouds have been detected on Uranus, perhaps related to seasonal changes in solar heating. The clouds are arranged in bands that circle the planet's rotation axis, running at constant latitudes parallel to the equator like the more vivid bands seen at Jupiter.

The features at different latitudes on Uranus move in the same direction as the planet rotates, but at faster speeds. The difference is greatest at high latitudes, where the clouds circle the poles in 14 hours, and it gets progressively smaller toward the equator, closer to the internal rotation period of 17.24 hours.

Since the high-latitude clouds are rotating faster than the interior of Uranus, the clouds cannot be simply carried by the planet's rotation. They are being blown by winds in the same direction as the planet rotates, just as clouds on Earth, Jupiter and Saturn. But unlike these planets, the winds on Uranus blow fastest near the poles of the planet.

The long, alternating periods of sunlight and darkness have little effect on the winds of Uranus. During the *Voyager 2* encounter, the south pole of Uranus was facing the Sun almost directly. The equator was in constant twilight, and the north pole had been in darkness for 20 years. So you might expect the south pole to be the warmest place on the planet and the north pole the coldest, with a temperature difference that would drive winds from pole to pole. But the thin clouds on Uranus move parallel to the equator, orthogonal to the direction of the pole.

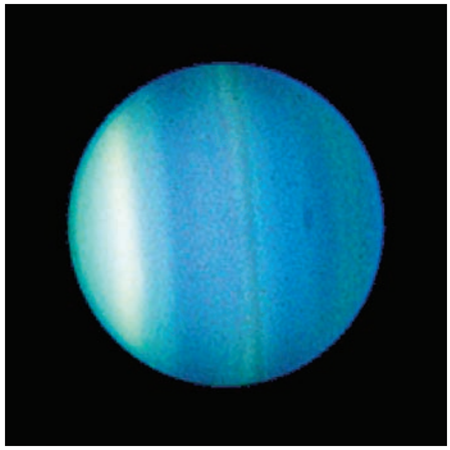


Fig. 11.2 Bands and vortex in the atmosphere of Uranus Although Uranus is similar in size and atmosphere composition to Neptune, it does not appear to have as active an atmosphere. A bright band at about 40 degrees southern latitude (*left*) nevertheless stretches out in the direction of rotation, and a dark, elongated whirling vortex, up to 3000 kilometers across, formed at about 27 degrees latitude in the northern hemisphere (*right*), which is just beginning to be exposed to sunlight after many years of being in shadow. This three-wavelength composite image was taken from the *Hubble Space Telescope* on 23 August 2006. (Courtesy of NASA/ESA/L. Sromovsky and P. Fry, U. Wisconsin, H. Hammel, STScI, and K. Rages, SETI Institute.)

The *Voyager* thermometer showed that the atmospheric temperatures are much the same everywhere on the globe. The temperature at the equator and the pole facing the Sun were about the same, and the dark winter side may have even been a few degrees hotter. So something has to be redistributing the solar heat, exchanging it between warm and cold places.

Stormy weather on Neptune

The *Voyager 2* flyby in 1989 forever changed our view of Neptune's weather. Despite its great distance from the Sun, the dimly lit atmosphere of Neptune is one of the most turbulent in the solar system, with violent winds, large dark storms and high-altitude white clouds that come and go at different places and times (Fig. 11.3).

Neptune has strong zonal winds driven and defined by the planet's rotation. The clouds near the equator circulate slower than Neptune's interior rotates so the prevailing winds blow in the westward direction, opposite to the

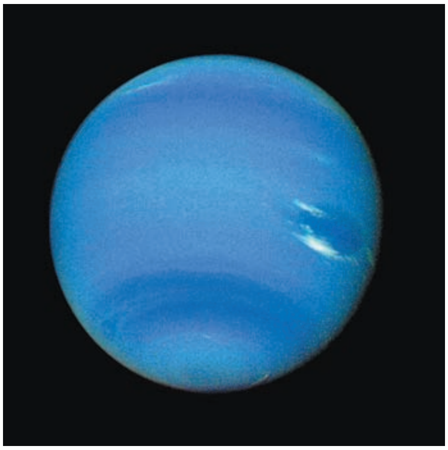


Fig. 11.3 Neptune's dynamic atmosphere White clouds lie above the swirling Great Dark Spot (*right center*), and raging winds reach speeds of 300 meters per second, creating a global banding in the atmosphere of Neptune. The faint sunlight at Neptune's great distance cannot provide the energy of such winds; they are probably energized by heat from the interior of the planet. This image was sent from the *Voyager 2* spacecraft on 14 August 1989, after a 12-year journey to the planet, using a 20-watt transmitter with less power than an ordinary incandescent light bulb. Traveling at the speed of light, the signals took more than 4 hours to reach Earth. (Courtesy of NASA/JPL.)

rotation of the planet. The equatorial wind speed on Neptune is about 325 meters per second relative to the core, almost as fast as Saturn's equatorial wind and faster than those on Jupiter or Uranus.

The wind pattern on Neptune lacks Jupiter's multiple zonal winds that flow in opposite directions. Neptune has just one westward air current at low latitudes, like the Earth's trade winds, and one meandering eastward current at mid-latitudes in each hemisphere, resembling the Earth's jet streams. And like Uranus, the polar and equatorial temperatures on Neptune are nearly equal.

The largest dark storms on Neptune are probably high-pressure systems that come and go with atmospheric circulation. The most prominent one was the Great Dark Spot, a vast, circulating storm almost as large as Earth. It is called the Great Dark Spot because it resembled the Great Red Spot of Jupiter and is dark rather than red. Both storms are found in the planetary tropics – at about one-quarter of the way from the equator to the south pole, both rotate counter-clockwise, in the direction of high-pressure anticyclones, and both are about the same size relative to their planet. As on Jupiter, some of the small dark spots on Neptune may be whirling in the opposite direction to the bigger one, perhaps indicating that they are little cyclones with descending material at their centers.

There are some important differences between the two Great Spots. Jupiter's red spot has survived for centuries, while Neptune's dark one disappeared from view within a few years of its sighting from *Voyager 2*. And Jupiter's storm lies above the clouds while Neptune's seems to form a deep well in the atmosphere, providing a window-like opening to the deeper, darker clouds below.

White, fleecy cirrus-like clouds cast shadows on the blue cloud deck below, indicating that they are high-altitude condensation clouds that rise about 100 kilometers above the surrounding ones. They form as atmospheric gas flows up, over and around the storm center, without being consumed by it. When the rising methane gas cools, it forms white clouds, fashioned from crystals of frozen methane. Water in the Earth's atmosphere freezes in a similar way into ice crystals that form white cirrus clouds. When strong upwelling carries the wispy white methane clouds to great heights in Neptune's atmosphere, they are sheared out like anvils of terrestrial thunderstorms.

Although the global wind pattern on Neptune resembles the Earth's trade winds and jet streams, they cannot be energized in the same way. Solar heating of the atmosphere and oceans drives the terrestrial winds. At Neptune's distance the Sun is 900 times dimmer and the winds should be correspondingly weaker if they are driven by the feeble sunlight. The fast winds on Neptune and the planet's complex stormy weather must instead be energized by heat generated in the planet's interior. The internal heat warms Neptune from the inside out, producing convection currents of rising and falling material, somewhat like a pot of boiling water on a stove. Uranus, on the other hand, shows no signs of substantial internal heating, and this may explain why its atmosphere is relatively benign and inactive.

You wouldn't want to forecast the variable and unpredictable weather on Neptune. When the *Hubble Space Telescope* (*HST*) took a look at the planet, the violent storms seen by the *Voyager 2* cameras had vanished without a trace, and other storms had appeared. The infrared detector on *HST* showed that Neptune's dynamic weather changes from day to day and over the years as well. White methane storm clouds as large as the Earth became more frequent in the southern hemisphere as it turned slowly toward the Sun's faint heat from local winter to spring (Fig. 11.4). In 2007, a team of astronomers using the Very Large Telescope in Chile reported that the temperature at Neptune's south pole had become 10 kelvin hotter than the rest of the planet, whose average temperature was about 73 kelvin. The extra heat in the warm south polar region provided an avenue for previously frozen methane to turn to gas and escape outward to colder cloud heights.

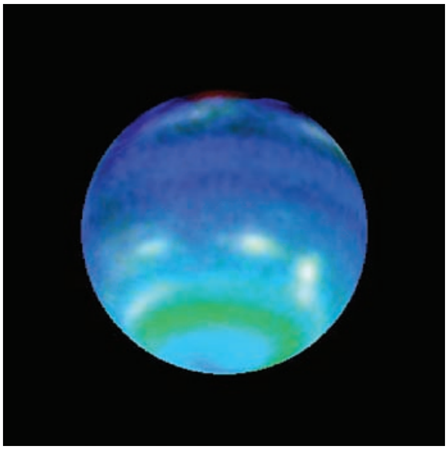


Fig. 11.4 Neptune's stormy disposition The weather on Neptune is recorded in this image, which combines simultaneous observations on 11 August 1998 with the *Hubble Space Telescope* and NASA's InfraRed Telescope Facility on Mauna Kea, Hawaii. The predominant blue color of the planet is a result of absorption of red and infrared light by Neptune's methane atmosphere. Clouds elevated above most of the methane absorption appear white. Neptune's powerful jet stream, where winds blow at nearly 400 meters per second, is centered at the dark blue belt near Neptune's equator. The Great Dark Spot detected when the *Voyager 2* spacecraft visited Neptune in August 1989 has completely disappeared in this image, taken nine years later. (Courtesy of NASA/STScI/IRTF.)

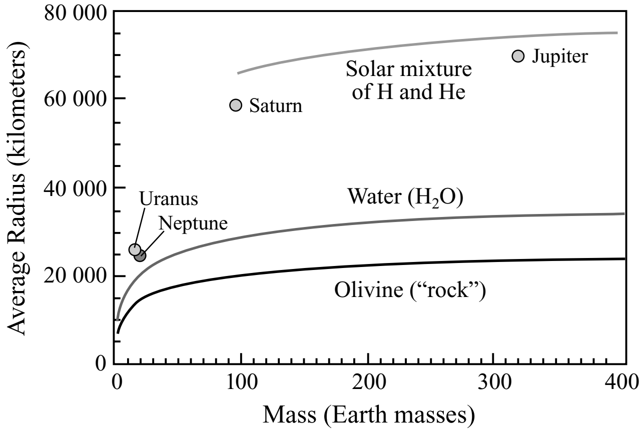


Fig. 11.5 Radius–mass relations A liquid body of solar composition describes a radius–mass relation (*top*) that approximates the mass and radius of Jupiter and Saturn. They consist mainly of the lightest element hydrogen (H) and next lightest abundant element helium (He). Uranus and Neptune contain little hydrogen or helium, for their radii are much too small to be consistent with a solar composition. Instead, they lie only slightly above the mass–radius relation for liquid water, so they probably contain large quantities of water. For comparison purposes, the bottom curve describes a giant planet composed of pure rock, as the Earth is but with a much smaller size.

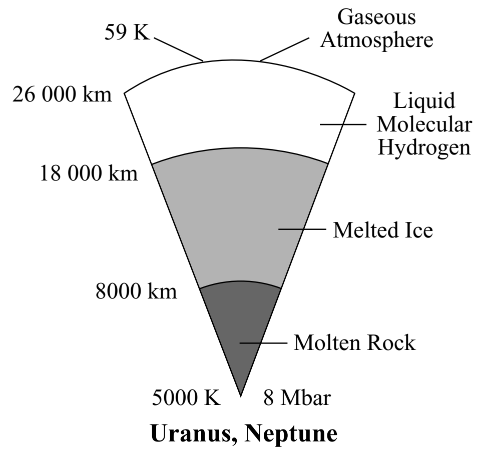


Fig. 11.6 Inside Uranus and Neptune An outer shell of liquid molecular hydrogen covers a thick inner shell of melted ice with liquid water, methane, and ammonia within the interior of both Uranus and Neptune. Because of their relatively low mass and hydrogen abundance, neither planet contains an inner shell of liquid metallic hydrogen, as Jupiter and Saturn do.

11.3 Interiors and magnetic fields of Uranus and Neptune

Uranus and Neptune are water worlds

Although Uranus and Neptune contain deep outer atmospheres of molecular hydrogen, their relatively small size, low mass, and high bulk mass density imply that hydrogen cannot be their main ingredient, unlike Jupiter and Saturn (Fig. 11.5). The hydrogen in Uranus and Neptune is confined within thin gaseous atmospheres and liquid molecular shells that do not extend to great depths and contribute only about 15 percent of the planetary mass (Fig. 11.6). These two planets do not have enough hydrogen, or sufficient mass and internal pressure, to squeeze the hydrogen into a metallic state. So there is no internal shell of liquid metallic hydrogen inside Uranus and Neptune.

The relatively large bulk density of Uranus and Neptune, when compared to Jupiter and Saturn, implies that heavy material contributes a much greater fraction of the mass for the two outermost major planets. Most of the interiors of Uranus and Neptune probably consist of a vast internal ocean of water, mixed with liquid methane and liquid ammonia (Fig. 11.6). So these planets are sometimes called water worlds, or alternatively ice giants since the liquids can all be frozen into ice at the colder cloud tops. These substances are nevertheless kept liquid inside the planets due to the higher temperatures there. Uranus and Neptune most likely also have a central molten rocky core beneath their oceans of melted ice.

So most of the interiors of Uranus and Neptune are not unlike the cores of Jupiter and Saturn, which are thought to contain 10 to 20 Earth masses of melted ice and molten rock.

The differences between the four giant planets apparently derive primarily from the amounts of hydrogen and helium that they were able to attract and hold on to as they formed. According to one explanation, the cores of Jupiter and Saturn accreted, or gravitationally gathered in, the surrounding hydrogen before it dissipated. The hydrogen dispersed into a larger volume and lower density at the more distant orbits of Uranus and Neptune. So their cores of ice and rock accumulated slowly and took a longer time to grow. Little hydrogen or helium was left to capture by the time they had grown large enough to start gravitationally collecting the surrounding gas. In another scenario, a blast of radiation from a luminous nearby star, other than the Sun, boiled away any hydrogen or helium that may have collected around Uranus and Neptune, while Jupiter and Saturn were protected by the greater density of the nearby material.

Tilted magnetic fields

Like the Earth, Jupiter and Saturn, both Uranus and Neptune have strong internal magnetic fields. But the resemblance ends there. Here on Earth our magnetic poles are very near our geographic poles, which is very useful for navigation with a compass. The magnetic and rotational axes of Jupiter and Saturn are also closely aligned. But they are way off kilter on both Uranus and Neptune.

The magnetic axis is tipped by 58.6 degrees to the rotation axis of Uranus, and the two are separated by 46.8 degrees for Neptune. By way of comparison, the displacement is 11.7 degrees on Earth, 9.6 degrees on Jupiter, and less than 1 degree on Saturn. If the disparity on Uranus and Neptune existed on Earth, a compass needle would point somewhere near Cairo, Egypt, instead of the North Pole. Theoreticians expected a closer alignment between rotation and magnetism on Uranus and Neptune.

It is almost certain that the same dynamo process as that responsible for Earth's magnetic field sustains the magnetic fields of Uranus and Neptune. This happens in Earth's molten metallic core, and it occurs within the liquid metallic hydrogen inside Jupiter and Saturn. Unlike these two giants, there is no shell of metallic hydrogen inside Uranus and Neptune. It is probable that the electrical conductivity within Uranus and Neptune is provided by water-rich material that has a conductivity that is about two orders of magnitude less than that of metals. It is also likely that this conductivity comes from protons, not electrons, within the ionized waters.

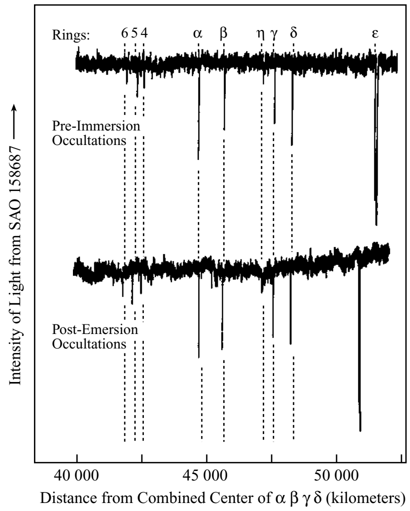


Fig. 11.7 Discovery of the thin rings of Uranus Astronomers recording the light of a star that was expected to disappear behind Uranus, on 10 March 1977, unexpectedly recorded short dips in the starlight before the star passed behind the planet (*top*). The same pattern was repeated when the star reappeared (*bottom*), indicating that narrow rings briefly block out the starlight at the same distance on opposite sides of the planet. The strong and abrupt absorption of starlight indicates that the narrow rings are quite opaque and have well-defined edges. These observations were taken from high above the Indian Ocean aboard the Kuiper Airborne Observatory. (Courtesy of James L. Elliot.)

11.4 Rings of Uranus and Neptune

Narrow, widely spaced rings around Uranus

Astronomers have had a history of happy accidents concerning Uranus, starting with William Herschel's (1738–1822) serendipitous discovery of the planet in 1781. Another lucky incident occurred on 10 March 1977, when the planet was scheduled to pass in front of a faint star. By observing such a stellar occultation, astronomers hoped to determine properties of the planet's atmosphere, and to accurately establish its size from the duration of the star's disappearance behind it.

Because of uncertainties in the predicted time of the star's disappearance, one telescope was set into action about 45 minutes early. Soon after the recording began, the starlight abruptly dimmed but then it almost immediately returned to normal, producing a brief dip in the recorded signal. At first, the dip was attributed to a wisp of cloud on Earth or to an unexpected change in the telescope's orientation. But the star blinked on and off several times before and after the planet covered it. Moreover, each dip on one side of Uranus was matched by another one on the other side, at the same distance from the planet (Fig. 11.7). The brief dips were due to something very

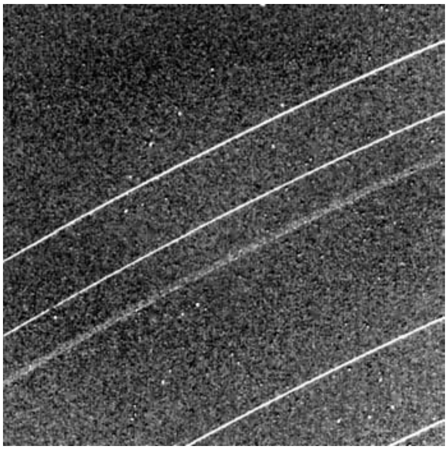


Fig. 11.8 The rings of Uranus This *Voyager* 2 image, taken on 23 January 1986, shows five of the nine rings of Uranus that had been previously inferred from Earth-based observations of their brief occultation of a star's light. In this view, sunlight striking the rings' particles was reflected back toward the camera, showing that the dense parts of the ring system consist of narrow rings with wide gaps. In contrast, Saturn's main rings are broad with narrow gaps. From bottom to top, the rings are designated by the Greek letters α, β, η, γ and δ. (Courtesy of NASA/JPL.)

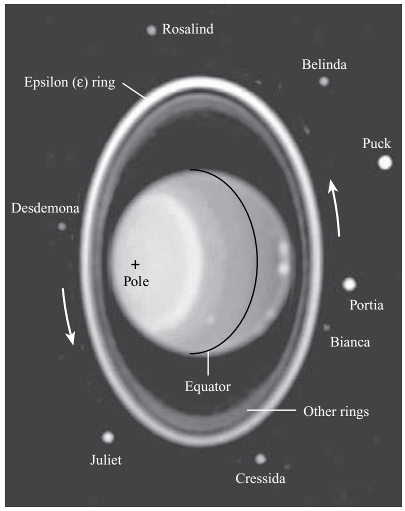


Fig. 11.9 Rings and small satellites of Uranus Eight of Uranus' small satellites circle the planet just outside its bright epsilon ring. This image, taken with the *Hubble Space Telescope* on 28 July 1997, is a false-color composite of three images taken at different infrared wavelengths in which Uranus appears relatively dim but the rings and moons do not. The satellites range in size from 40 kilometers across for Bianca, to 150 kilometers for Puck. The arrows denote their direction of revolution about Uranus. White clouds are seen just above the planet's blue-green methane atmosphere. (Courtesy of NASA/STScI.)

narrow and very close to Uranus, and their mirror-like symmetry indicated that a family of narrow rings, which blocked out the star's light but could not be seen directly from the Earth, surrounds the planet.

During the next few years, observations of more than 200 stellar occultations by Uranus revealed the details of nine narrow rings. In order of increasing distance from Uranus, the rings are named 6, 5, 4, α, β, η, γ, δ and ε, following the differing notation of the discoverers. From the brief duration of the dips of blocked starlight, astronomers concluded that all but one of the individual rings could be no wider than 10 kilometers. The relatively long time between the dips indicated that the threadlike rings are separated by hundreds of kilometers of nearly empty space. These skeletal, web-like rings are unlike any seen before, all exceptionally narrow and widely spaced from each other. Since the rings are so narrow, and separated by wide spaces of almost nothing, they are extremely difficult to see with even a large telescope on Earth.

When *Voyager 2* arrived at Uranus in 1986, nearly a decade after the discovery of its narrow rings, instruments on the spacecraft confirmed all the known rings, and added at least two (Fig. 11.8). They found the λ ring, a narrow strand between the δ and ε rings, and another one interior to ring 6. The spacecraft also discovered at least 10 small moons that are located just outside the ring system.

The irregular orientation, narrowness, and shapes of the rings of Uranus are attributed to nearby small moons (Fig. 11.9). The repeated gravitational tugs of two of them, Cordelia and Ophelia, pull the ε ring into its oval shape and restrain its edges. These tiny moons flank the ring, controlling its shape in much the same way that the shepherd satellites, Pandora and Prometheus, constrain Saturn's F ring. Nearby moons probably sharpen the edges of the other rings, keeping them from spreading out as the result of particle collisions.

The *Hubble Space Telescope* was used to discover two fainter outer rings of Uranus, which were not fully recognized until 2005. That brought the total of known dusty rings of Uranus to 13. The space telescope has also been used to image the rings as they slowly closed up, as viewed from near Earth, snagging a rare edge-on view in August 2007 (Fig. 11.10). Astronomers only see the ring edges every 42 years as the planet follows its leisurely 84-year orbit about the Sun.

Both the ring particles and the inner moons of Uranus are very dark, quite unlike the bright particles and tiny moons found in Saturn's wide rings. In fact, the particles of Uranus's rings reflect only about 2 percent of the sunlight falling on them, making them as dark as charcoal. Most investigators agree that the material is dark because it is rich in carbon, either derived from a methane coating or primordial in origin.

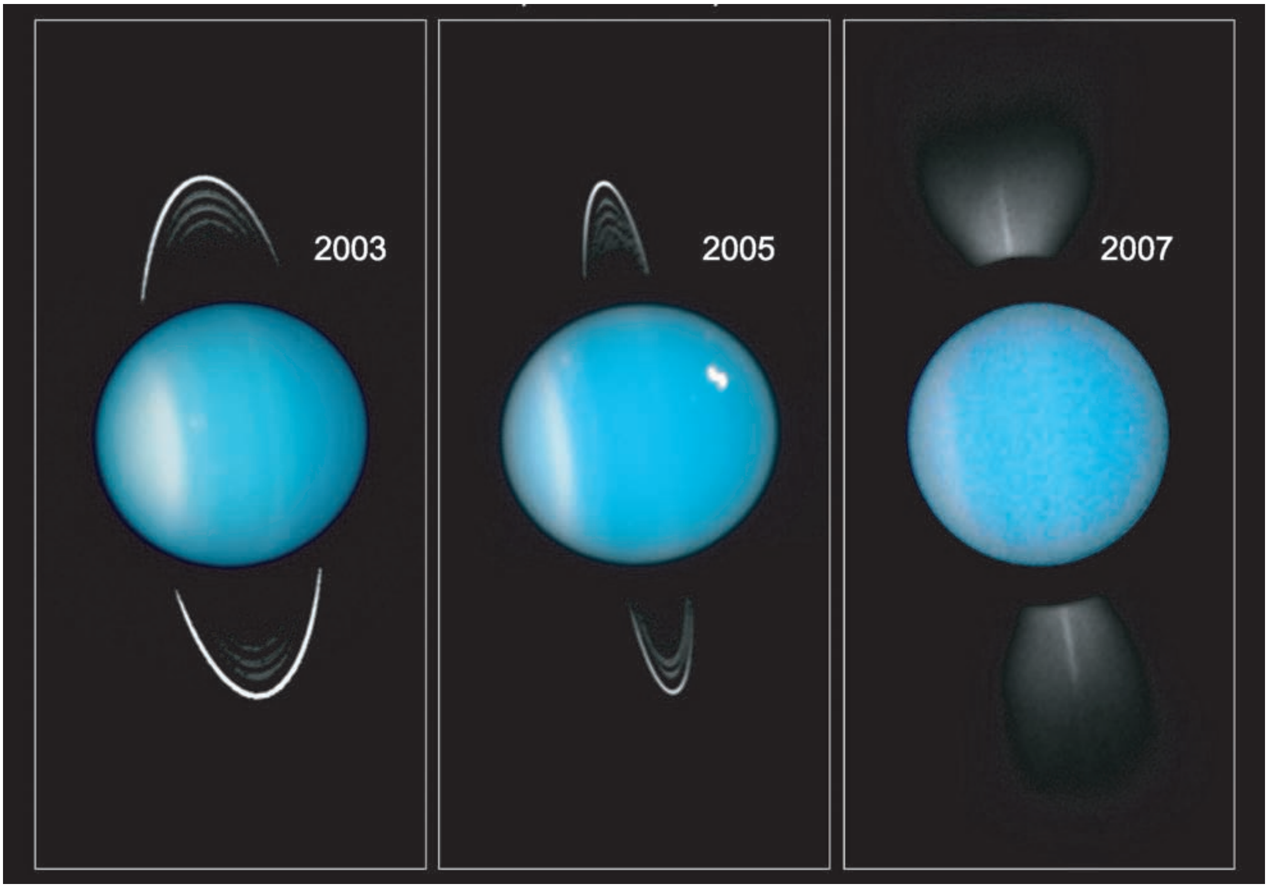


Fig. 11.10 Rings on edge These images were taken from the *Hubble Space Telescope* over a four-year period, from 2003 to 2007 (*left to right*), as Uranus moved along one-twentieth of its 84-year journey around the Sun. As viewed from the Earth, the rings are always nearly perpendicular to the planet's orbital plane, but they gradually close up from partly open to edge on. When tiled edge-on to Earth, on 14 August 2007, the rings appeared as spikes above and below the planet, within a fan-shaped glare – an image artifact. The fainter outer rings appear in the 2003 image, but were not noticed until they were seen in the 2005 image. Uranus has a total of 13 dusty rings. (Courtesy of NASA/ESA/M. Showalter, SETI Institute.)

Neptune's sparse thin rings and arcs

After the discovery of the rings of Uranus by watching a distant star pass behind the planet, astronomers hoped to repeat the achievement by observing stellar occultations by Neptune, but the results were inconclusive. Sometimes the starlight would remain unchanged before and after the planet directly occulted the star. At other times the star would blink on and off, but always on just one side of the planet. Because the brief dimming of starlight was not symmetrical about the planet, and not all stellar occultations produced a blinking signal, the hypothetical rings became shortened, in the minds of the astronomers, to ring-arcs that only reached partway around the planet. Chance might then dictate which astronomers would detect the obscuration.

Voyager 2 clarified the problem. Neptune's ring-arcs turned out not to be isolated segments, but rather three thicker portions of one very thin ring (Fig. 11.11). The ring is continuous, stretching all the way round the planet just like any well-behaved ring, but it does not have an even width or density all around. Its material is generally spread so thinly that it does not noticeably dim a star's light, and the ring is only wide and dense enough to hide a star in three arc-like concentrations. They have been named Liberté, Egalité and Fraternité after the French revolutionary slogan. It was these wide high-density clumps that had been detected from Earth, blocking starlight and giving the impression of disconnected arcs. The rest of the ring couldn't be seen from Earth because it is so transparent, and hence below the threshold of detection.

The existence of such clumps, or concentrations, in the rings was an enigma. Every time that the ring particles collide and bump together, they must change speed, and gradually spread around the ring away from the clumps.

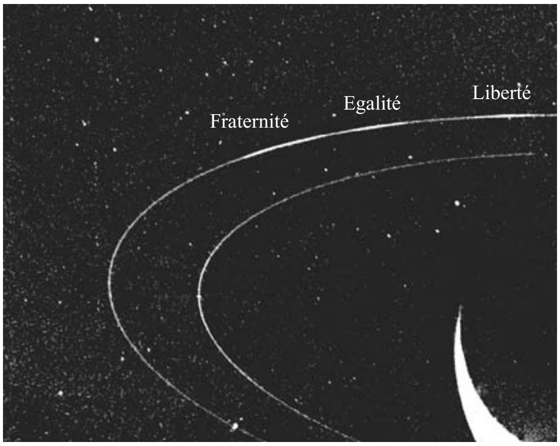


Fig. 11.11 Neptune's rings As *Voyager 2* left Neptune in August 1989, the planet's narrow rings were backlit by the Sun, enhancing the visibility of the rings' dusty particles. The outer ring consists of at least three dense clumps of orbiting debris, named Liberté, Egalité and Fraternité, which stand out from the thinner remainder of the ring. Astronomers on the ground had only detected the clumps during some stellar occultations, and assumed that the ring was incomplete. The outermost ring is named Adams, and the innermost is designated Le Verrier. They are named for John Couch Adams (1819–1892) and Urbain Jean Joseph Le Verrier (1811–1877) who independently predicted the existence of the then unknown planet Neptune. A third, fainter ring is located closer to Neptune than the two main rings; it has been named Galle after Johann Gottfried Galle (1812–1910) who found the planet close to both of Adams' and Le Verrier's predicted positions using a 0.23-meter (9-inch) refractor. (Courtesy of NASA/JPL.)

Unless something is confining the material in the arc-like concentrations, it should spread uniformly around the entire ring in just a few years. Some external force must therefore be holding the material in place, and keeping it within the three arcs. A small moon might hold the clumps together by its gravity, and astronomers think they have found it – the satellite Galatea orbits Neptune just inside the ring.

Altogether, *Voyager 2*'s cameras found five rings around Neptune, named in honor of astronomers. From the innermost to outermost, they are the Galle, Le Verrier, Lassell, Arago and Adams rings. The faint, innermost Galle ring is broad, roughly 1700 kilometers wide, and may extend all the way down to the top of the planet. The outermost Adams ring is bright and narrow, and contains the three bulging clumps.

Unlike the main narrow rings of Uranus, the brightest Adams and Le Verrier rings of Neptune contain a vast amount of microscopic dust, apparently produced by grinding down or eroding larger particles. In fact, the sparse, dusty rings of Neptune contain only about a thousandth as much matter as the rings around Uranus, and a million times less matter than Saturn's rings. Put together, all the particles in Neptune's rings would make a body only a few kilometers across. So small nearby satellites could have been the source of Neptune's rings. Meteoritic bombardment over the eons could, for example, have pulverized such a moon into rubble, producing all that dust. Even now, three small moons are located deep within Neptune's ring system, between the Galle and Le Verrier rings.

Formation and evolution of the rings of Uranus and Neptune

The austere rings that now circle Uranus and Neptune may have had a violent and chaotic past, arising from catastrophic collisions of moons or when one larger satellite moved inward by tidal interaction with the planet until it was close enough to be ripped into pieces. The inner moons and larger particles in the rings were then probably gradually broken up by collisions into smaller ones.

All the ring particles we see today might eventually be eroded away by meteoritic bombardment, ground into fine dust by particle collisions, or displaced by gravitational interaction with neighboring satellites. Once the rings have been turned into dust, as they have in parts of Uranus's ring system and most of Neptune's, they must eventually spiral inward into the planet's atmosphere where they will burn up. Thus an entire dusty ring system will inevitably decay and disappear over astronomical times.

This need not imply that rings will vanish from Uranus and Neptune. The rings could easily be replaced by the collisional breakup or meteoritic erosion of tiny moons already embedded in the rings. If you broke up all the satellites now within Neptune's rings, you might produce rings as magnificent as Saturn's present ones. This now brings us to some of the larger satellites lying outside the rings of these two planets.

11.5 The large moons of Uranus and Neptune

Five major moons of Uranus

Uranus possesses five major satellites discovered telescopically from Earth before the space age, and named Miranda, Ariel, Umbriel, Titania and Oberon (Table 11.2), and as the result of modern telescopes and the *Voyager 2* spacecraft we now know that the planet possesses at least 27 moons. All of the large satellites orbit Uranus near its equatorial plane, perpendicular to its orbital plane, and they are

Table 11.2 The five large moons of Uranus[a]

Name	Mean distance from Uranus (radii)	Orbital period (days)	Radius (kilometers)	Mass (10^{21} kilograms)	Mass density (kg m^{-3})
Miranda	5.08	1.41	236	0.065	1214
Ariel	7.47	2.52	579	1.29	1592
Umbriel	10.41	4.15	585	1.22	1459
Titania	17.07	8.70	789	3.42	1662
Oberon	22.82	13.46	761	2.88	1559

[a] The orbital distances are given in units of the radius of Uranus, which is 25 559 kilometers, or about four Earth radii. The satellite radii are given in units of kilometers. By way of comparison, the mean radius of the Earth's Moon is 1738 kilometers. The mass is given in units of 10^{21} kilograms – our Moon's mass is 73.5 in these units. The mass density is given in units of kilograms per cubic meter (kg m^{-3}).

all tidally locked into synchronous orbits with one side perpetually facing the planet. As a group, they are similar in size to the mid-sized satellites of Saturn. The two largest and outermost moons, Titania and Oberon, are roughly half the size of the Earth's Moon; the smallest and innermost, Miranda, is about one-seventh the lunar size. Infrared spectroscopy from Earth indicated that they all have water ice on their surfaces, but their icy surfaces are dark looking. Their bulk mass densities (Table 11.2) imply that the four largest moons of Uranus are about half rock and half water ice, so they are rocky on the inside, as well as dirty on the outside.

Because they are relatively small and very cold outside, these moons were expected to be heavily cratered ice-balls, devoid of any signs of internal activity. When *Voyager 2* photographed their icy surfaces at close range, the expected craters were seen, but the surfaces also displayed surprising diversity. They have been fractured and cracked open, probably as the result of surface expansion, and covered by flowing ice or perhaps even liquid water that subsequently froze. Internal heat supplied by radioactive elements may have produced the surface expansion, melted the ice, and generated icy volcanic flows. Because radioactive elements are embedded in rock, the relatively large amount of rock in the large moons of Uranus, as compared to Saturn's satellites of about the same size, produces greater internal heat.

It was Prospero's daughter Miranda who, in Shakespeare's *The Tempest*, proclaimed:

> O, Wonder!
> How many goodly creatures are there here!
> How beauteous mankind is! O, brave new world,
> That has such people in't!
>
> *The Tempest*, Act V, Scene i, lines 182–5

and her counterpart at Uranus certainly has all the earmarks of a "brave new world".

The landscape on Miranda is one of the most amazing yet observed in the solar system. It includes deep canyons, old cratered plains, bright younger terrain, and an eclectic mixture of ridges, grooves, mountains, valleys, fractures and faults (Fig. 11.12).

There are two possible explanations for Miranda's jumbled surface. One explanation supposes that the satellite was once blasted apart by one or more catastrophic collisions and then pulled together again under the combined gravitation of its pieces, reforming into a born-again moon. It may even have been broken up and reassembled several times. The alternative explanation holds that Miranda is a half-grown world whose development was interrupted. It supposes that the moon heated up inside, soon after formation, and that the dense rocky material began to sink toward the center while the lighter icy substances started rising to the surface. But as Miranda cooled the internal heat dwindled, and the differentiation process could not be finished.

Neptune's Triton, a large moon with a retrograde orbit

Jupiter, Saturn and Uranus have a flock of satellites whose orbits mimic those of the planets around the Sun. Their larger moons revolve in regularly spaced, circular orbits in the same direction as the rotation of the planet and close to the planet's equatorial plane, presumably because they share the rotation of the nebular disk from which the planet and its satellites formed. The radii, orbital distances and other characteristics of these regular satellites also tend to differ in smooth progression. In sharp contrast to Jupiter,

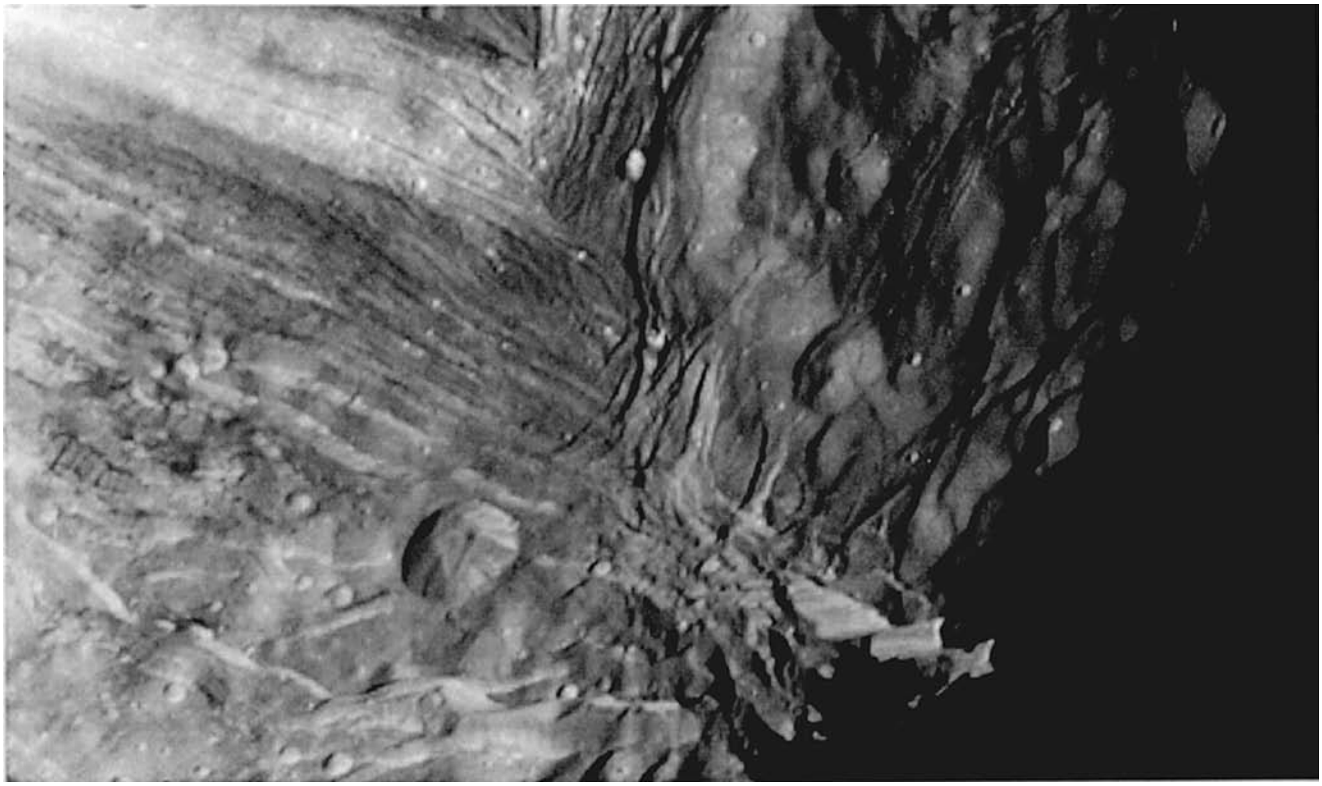


Fig. 11.12 Miranda The innermost large moon of Uranus, Miranda, is seen at close range in this *Voyager 2* image, which was taken on 24 January 1986. The great variety and directions of the fractures and troughs, and the different densities of impact craters on them, signify a long, violent history for Miranda. Two distinct types of terrain are visible. They are rugged, higher-elevation terrain, disappearing into the shadows (*upper right*), and the lower-elevation, striated terrain (*upper left*). The light gray cliff (*bottom right*) is 20 kilometers high; the larger crater (*lower middle*) is about 25 kilometers across. (Courtesy of NASA/JPL.)

Saturn and Uranus, the planet Neptune lacks a system of regular larger satellites.

Neptune's largest moon, Triton, is the only large satellite in the solar system to circle a planet in the retrograde direction, opposite to the planet's direction of rotation (Fig. 11.13). This oddity is compounded by the high orbital inclination. The satellite's orbital plane is tilted at an enormous 157 degrees from the planet's equator. Nereid, the outermost moon of Neptune, discovered by Gerard Kuiper (1905–1973) in 1949, adds to the mayhem. It has the most elongated orbit of any planetary satellite, seven times as distant from the planet at its farthest compared with its closest approach.

Triton's unusual retrograde orbit was quickly established, soon after its discovery – by William Lassell (1799–1880) in 1846, just three weeks after the discovery of Neptune. But due to its great distance, the radius, mass and reflectivity of Triton remained uncertain until 25 August 1989 when instruments on the *Voyager 2* spacecraft measured them. With a radius of about 1353 kilometers, Triton is about three-quarters of the radius of the Earth's Moon, at 1738 kilometers. The slight gravitational tug exerted on *Voyager 2* by Triton yielded a mass of 2.141×10^{22} kilograms for the satellite. The size and mass are combined to give a mean mass density of about 2061 kilograms per cubic meter. This is significantly more dense and rock-rich than the large moons of Uranus, between 1400 and 1700 kilograms per cubic meter.

Triton's frozen surface, thin atmosphere and geyser-like eruptions

Measurements from *Voyager 2* indicated that Triton is the ultimate icebox, with a daytime surface temperature of 37.5 kelvin. This deep freeze is only about three dozen degrees above absolute zero, when nothing can move, not even an atom. In fact, Triton has the coldest measured surface of any natural body in the solar system! It is so cold because it is very far away from the Sun, therefore receiving little sunlight, and also because Triton reflects more of the incident sunlight than most satellites – only Enceladus and Europa are comparable. As a result, the total amount of sunlight absorbed by Triton's surface is less than that of any other planet or satellite. A frosty ice coating overlies

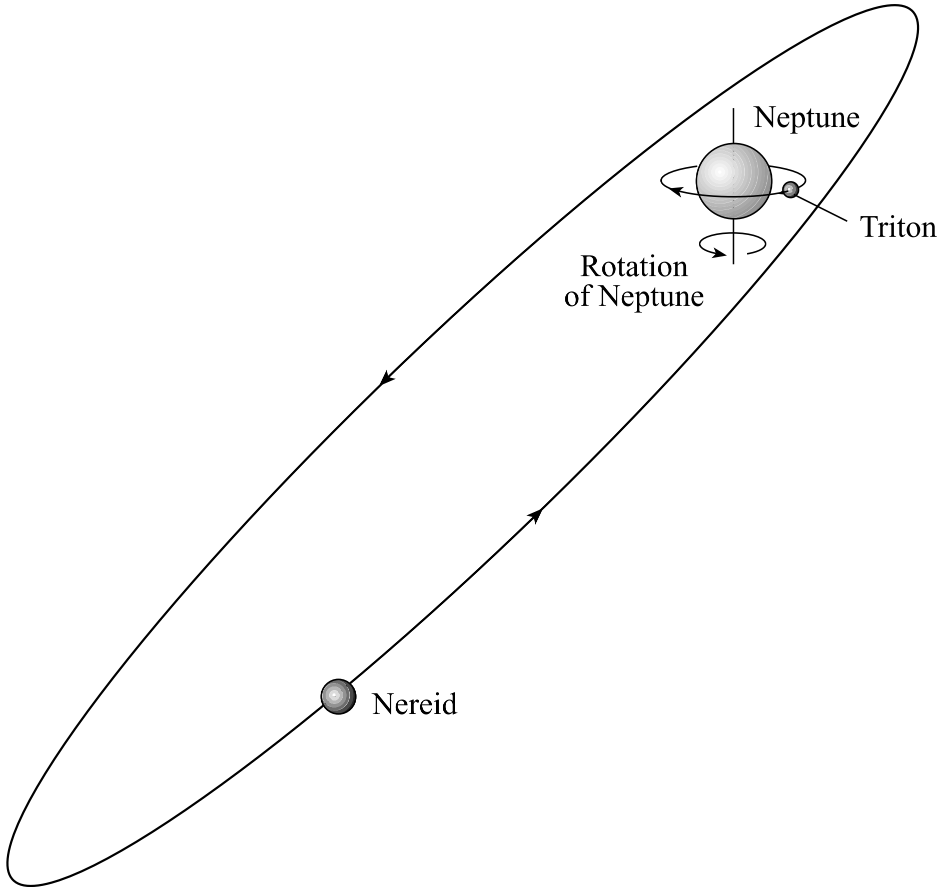


Fig. 11.13 Neptune's odd satellites Nereid, the outermost of Neptune's satellites, travels in a highly inclined, eccentric orbit, in the same direction as that of the planet's rotation. Triton, the largest satellite of Neptune, travels around Neptune in a circular obit, but, unlike any other largest satellite of the giant planets, it travels in the opposite retrograde direction to the rotation of Neptune. In addition, careful analysis of Triton's motions shows that the satellite is in a decaying orbit and is slowly being pulled toward Neptune.

all of the surface features, reflecting the incident sunlight (Fig. 11.14). The brilliant ice has a salmon-pink tint with peach hues, possibly due to organic compounds derived from methane by the bombardment of energetic particles from the solar wind and Neptune's radiation belts.

Although nitrogen and methane frosts are apparently the dominant constituents of Triton's visible disk, water ice is needed to support and preserve the observed topography, including cliffs and ridges that exceed one kilometer in height. At the frigid temperature of Triton, water ice is as strong as steel, and behaves like hard rock on Earth; but the methane and nitrogen ice do not have sufficient strength to support the elevated features, which would deform and collapse under their own weight. Thus, thin, brilliant veneers of nitrogen and methane ice apparently overlie a rigid crust of water ice.

Despite its frigid surface conditions, an exceedingly tenuous atmosphere envelops Triton. It consists mainly of nitrogen molecules, the same gas that makes up most of the atmospheres of Earth and Saturn's moon Titan. But the cold, thin atmosphere on Triton has a surface pressure of only 15 microbars, or 15×10^{-6} bars, which is ten-millionths the surface pressure of the atmosphere on Titan and fifteen-millionths of the air pressure at sea level on Earth.

Voyager 2's cameras showed that Triton's southern polar cap was then in the process of dissipating. Nitrogen ice was sublimating, or changing directly from the solid to vapor form, and supplying the atmosphere with gas. The vaporized nitrogen gas is probably carried by atmospheric winds to the dark northern hemisphere.

Even in the middle of southern summer, a bright ice cap extends from the south pole three-quarters of the way to Triton's equator (Fig. 11.14). It is so cold that some of Triton's air freezes out at its poles, coating them with a huge ice cap of frozen nitrogen. It is also too cold for water ice to vaporize from Triton's surface and enter its atmosphere. By way of comparison, the Earth's polar caps contain frozen water ice, for it is too warm for nitrogen to freeze at our planet's poles.

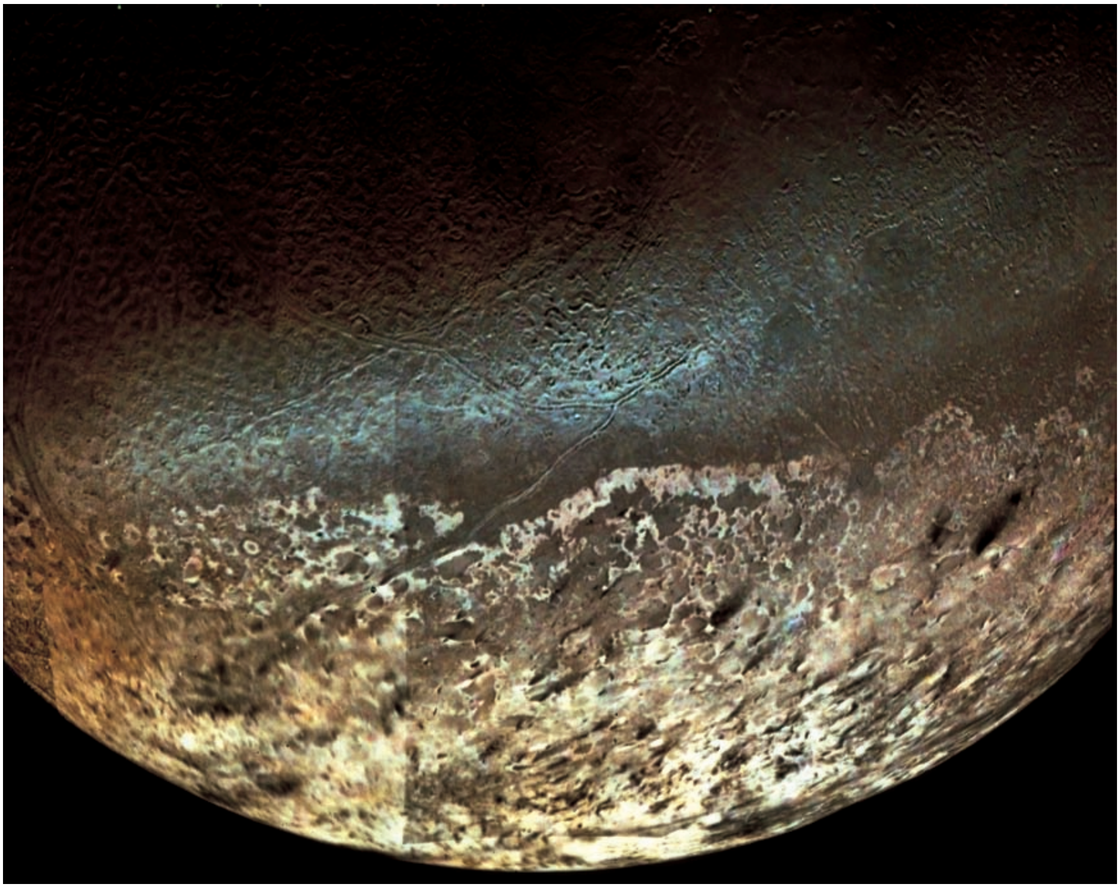


Fig. 11.14 Triton, the largest moon of Neptune A mosaic of the south polar cap of Triton, the largest satellite of Neptune, taken in August 1989 from the *Voyager 2* spacecraft. At the time of this flyby, Triton was the coldest measured object in the solar system with a surface temperate of 37.5 kelvin. It is so cold that most of Triton's thin nitrogen atmosphere is condensed as frost, making it the only satellite in the solar system known to have a surface coated with deposits of solid nitrogen ice. Small admixtures of highly reflective methane ice are also present on the surface; they have been reddened and darkened by the action of energetic radiation. Dark plumes and streaks probably mark vents of ice volcanoes, where nitrogen gas has been driven outward in geyser-like eruptions from beneath the surface and carried by the winds into elongated streaks. (Courtesy of NASA/JPL.)

Numerous parallel dark streaks in the bright southern cap are apparently blown by the prevailing winds in Triton's tenuous atmosphere and strewn over the ice. The dark streaks emanate abruptly from specific points, and thin away toward their ends. Some are only a few kilometers in length, while others extend more than 100 kilometers. They must have formed relatively recently, for they seem to overlie deeper ice deposits, and it is unlikely that they could survive sublimation of the polar ice.

The dark streaks are probably related to jets and geysers of nitrogen seen erupting from the sunlit polar ice. These plumes rise in straight columns to an altitude of about 8 kilometers, where dark clouds of material are left suspended and carried downwind for over 100 kilometers, like smoke wafted away from the top of a chimney. Since the active plumes occur where the Sun is directly overhead, the solar heat might energize them.

Scientists have not reached a consensus about what produces the plumes, but one likely explanation is that geysers are sending up plumes of nitrogen gas laced with extremely fine dark particles. Triton is far too cold to spout steam and water, like geysers on Earth, but Sun-powered geysers might expel dark material when pent-up nitrogen gas becomes warm and breaks through an overlying seal of ice.

On Triton, the subterranean heat might be accumulated from sunlight, which passes through the translucent ice and is absorbed by darker methane or other carbon-rich material encased beneath. The overlying nitrogen ice would trap the solar heat, for it is opaque to thermal infrared radiation, producing a solid-state greenhouse effect. Nitrogen gas, pressurized by the subsurface heat, then explosively blasts off the iced-over vents or lids, launching volcanic plumes of gaseous nitrogen and

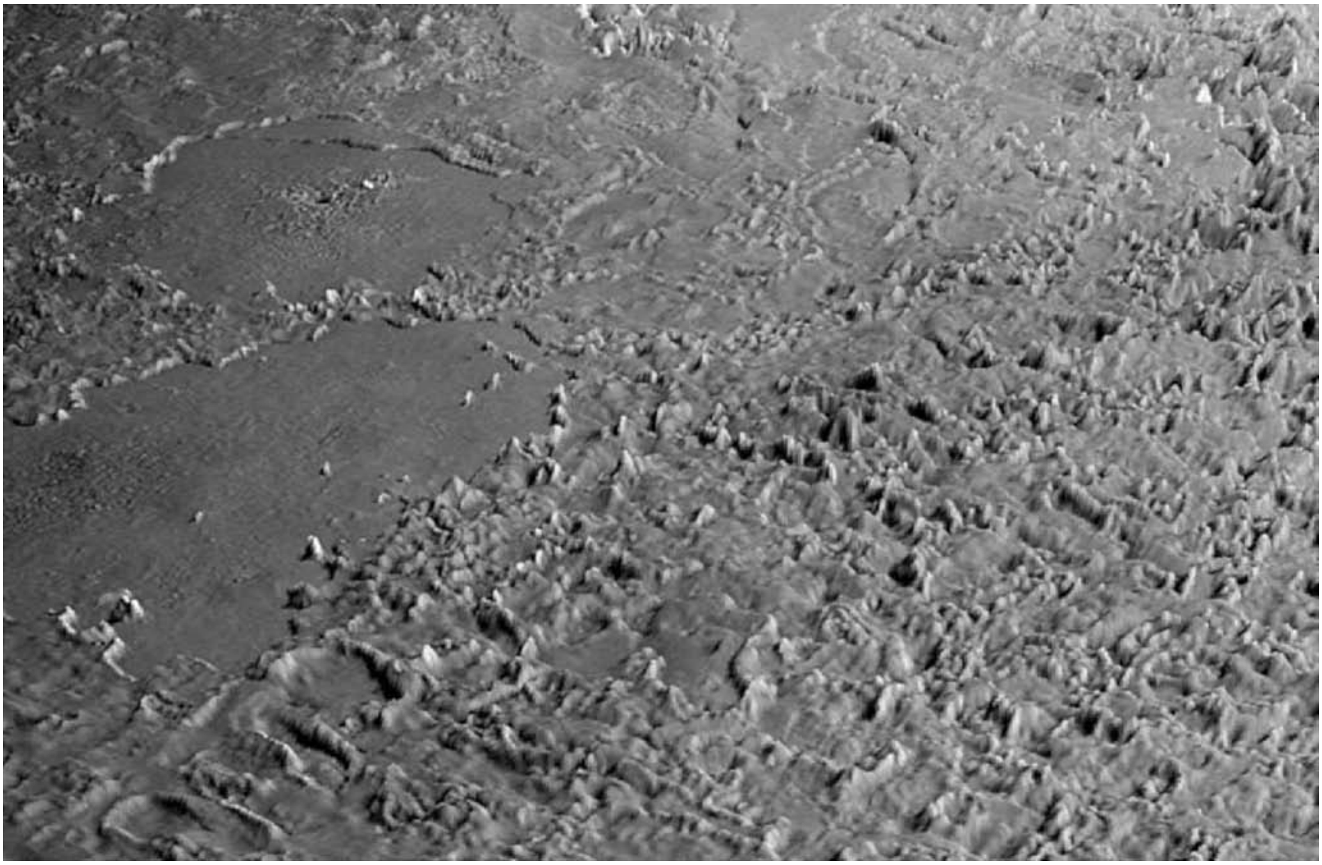


Fig. 11.15 Smooth and bumpy terrain on Triton Smooth volcanic plains, created by ice volcanoes on Triton, form the flat, frozen surfaces of ice lakes (*left*), most likely filled with water ice and perhaps coated with deposits of nitrogen ice. The absence of any impact craters indicates that the surface is relatively young. The rugged terrain (*right foreground*) is Triton's cantaloupe terrain, a network of interfering, closely spaced dimples or depressions termed cavi, each 25 to 35 kilometers across. They are also of internal origin, but not due to volcanic flow. The cantaloupe terrain may be explained by a gravitational instability in which less dense material rises through overlying dense material, overturning the icy crust. The rising blobs of ice are known as diapirs. (Courtesy of NSA/JPL.)

ice-entrained darker material into the tenuous atmosphere, just as the water in an overheated car radiator is explosively released when the radiator cap is removed.

Astronomers expected to find a surface covered by craters, but there are almost no craters in sight on Triton. Much of the visible surface outside the polar cap resembles the skin of a cantaloupe, containing numerous pits or dimples with low raised rims and ridges that snake their way through them (Fig. 11.15). But the dimpled pits are too similar in size and too regularly spaced to be impact craters, and appear to have an internal origin.

Long cracks or faults on Triton seem to have been partially filled with oozing water ice. Vast frozen basins found within Triton's equatorial regions have apparently been filled by icy extrusions flowing out from the warm interior, like a squeezed slush cone. These frozen lakes of water ice look like inactive volcanic calderas, complete with smooth filled centers, successive terraced flows and vents (Fig. 11.15).

Origin and evolution of Triton

Triton's retrograde and inclined orbit suggests it may not date back to the planet's formation, and that Triton was once a separate world which was captured into an eccentric, backwards and tipped orbit in the remote past. Neptune's lack of an ordered family of large satellites might also be explained if Triton was born in its own independent orbit around the Sun, and was subsequently captured by Neptune, destroying any regular satellite system the planet may have had in the process.

Neptune could only pull Triton into its gravitational sphere of influence if the passing body lost some energy as it went by. Otherwise, Neptune and Triton would just continue to go on their separate ways. One possibility is that Triton smashed into one of the planet's more substantial satellites, losing enough energy that it could not escape Neptune's gravitational pull and went into orbit around it. Or Triton might have grazed the fringes of Neptune's

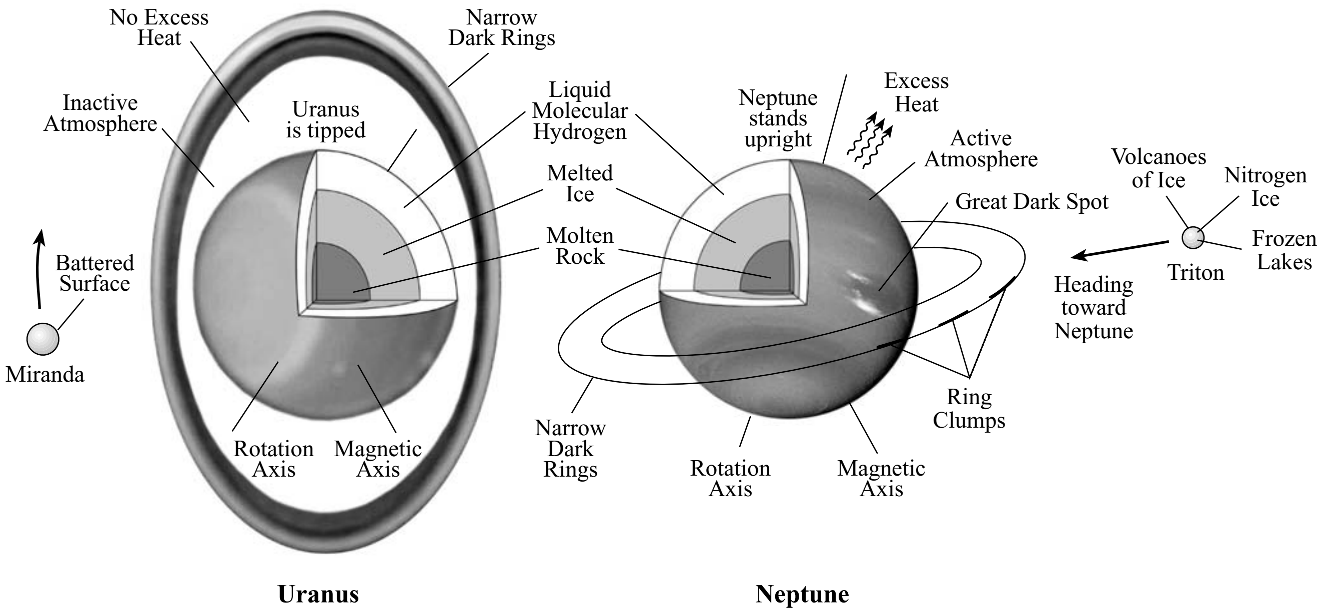


Fig. 11.16 Summary diagram

atmosphere where friction slowed it down enough to be captured.

When Triton began revolving around Neptune, its orbit would have most likely been very elliptical. But tides raised on Triton by Neptune could have circularized the satellite's orbit. While its orbit was evolving, Triton could have cannibalized or ejected satellites it encountered, thereby removing any large regular satellites Neptune may have once had, while also knocking Nereid into its unusual, highly eccentric orbit.

This explanation for Triton's origin is considered likely but still speculative, for we do not know for certain that the moon formed separately from Neptune. But we are fairly certain about Triton's suicidal future. Tidal interaction between Neptune and Triton is now gradually drawing the moon inward. Some time in the very distant future, Triton will be pulled close enough to Neptune for the planet's tidal forces to rip the moon apart, perhaps forming rings as bright and magnificent as Saturn's are today.

Part 4 Remnants of creation: small worlds in the solar system

12 Asteroids and meteorites

- There are billions of asteroids in the main asteroid belt, located between the orbits of Mars and Jupiter.
- The asteroid belt is largely empty space, and a spacecraft may safely travel through it.
- Hundreds of Trojan asteroids circle the Sun in the same orbit as Jupiter. These asteroids are located near the two Lagrangian points where the gravity of the Sun balances that of Jupiter.
- The Earth resides in a swarm of asteroids. Many of these near-Earth asteroids travel on orbits that intersect the Earth's orbit, with the possibility of an eventual devastating collision with our planet.
- Asteroids can be chaotically shuffled out of certain orbits in the main belt, and redirected into the inner solar system.
- The asteroids are the pulverized remnants of former worlds that failed to coalesce into a single planet.
- Groups of asteroids, known as families, have very similar orbits. The members of each family are the collision fragments of a larger object, which was itself much smaller than a major planet.
- The combined mass of billions of asteroids is less than 5 percent of the mass of the Earth's Moon.
- The largest body in the main asteroid belt, 1 Ceres, and the first to be discovered there, is about 950 kilometers across and contains about one-third of the total mass of all the asteroids.
- Unlike most, if not all, of the other asteroids, 1 Ceres has a round shape, suggesting a differentiated interior with a rocky core, thick mantle of water ice, and a dusty outer crust. 1 Ceres has been designated a dwarf planet.
- The gravity of the large majority of asteroids is too weak, and the mass too low, to pull them into a round shape. Their irregular, chipped shapes have been formed by eons of collisions.

- The colors of sunlight reflected from asteroids indicate that they formed under different conditions prevailing at varying distances from the Sun.
- Roughly 75 percent of the main-belt asteroids are the dark, black carbonaceous C-type orbiting the Sun in the outer half of the belt; about 15 percent of asteroids are bright, red, silicate S-type, residing on the sunward side of the main belt. Metallic, M, asteroids account for some of the remaining ones, and a few V asteroids are covered with volcanic basalt.
- Asteroids may be mined for minerals or water.
- Periodic brightness variations tell us that most asteroids are non-spherical objects spinning with periods of a few hours.
- Radar images of asteroids reveal diverse shapes and cratered surfaces, ranging from solid rock to loose rubble, as well as binary asteroids and even a triple one.
- The close-up view obtained by passing spacecraft indicates that asteroids have been battered and broken apart during catastrophic collisions in years gone by.
- Some asteroids are thought to be rubble piles, the collected fragments of past collisions held together by gravity; other asteroids are solid rocks of uniform internal composition. Asteroids 253 Mathilde and 25143 Itokawa are rubble piles, whereas asteroid 433 Eros is a battered but solid rock.
- The *Near Earth Asteroid Rendezvous* (*NEAR*) spacecraft was the first to orbit an asteroid and the first to land on one. *NEAR* circled the near-Earth asteroid 433 Eros for a year, examining its dusty, boulder-strewn landscape in great detail, obtaining an accurate mass for the asteroid, and showing that much of it is solid throughout.
- The Japanese *Hayabusa* spacecraft has orbited the small, near-Earth asteroid 25143 Itokawa, whose orbit crosses that of the Earth. This asteroid, just 535 meters in its longest dimension, has a lumpy, oblong shape composed of two parts with rough terrain joined by a smooth region. It is a loose collection of rubble held together by its weak gravity, with surface rocks sorted by shaking and jostling. *Hayabusa* landed on the asteroid and may return a sample of it to Earth.
- The *Dawn* spacecraft, launched in 2007, is expected to orbit the asteroid 4 Vesta from September 2011 to April 2012 and to orbit the dwarf planet–asteroid 1 Ceres in 2015. Astronauts may land on an asteroid by 2025.
- Asteroid rotation periods range from a couple of minutes to a few months, and some of them wobble instead of rotating uniformly.
- Meteorites are rocks from space that survive their descent to the ground.
- The number of recovered meteorites more than doubled when scientists discovered a large number of them on the blue ice of Antarctica.
- A few meteorites may have been blasted off the Moon or Mars, but most of them are chips off asteroids.
- The organic matter found in meteorites predates the origin of life on Earth by a billion years; but the meteoritic hydrocarbons are not of biological origin.

12.1 The orbits of asteroids

The main belt of asteroids

Millions of asteroids are confined within a wasteland between the orbits of Mars and Jupiter, like so much rubble left over from the creation of the solar system. Most of them occupy a great ring, known as the asteroid belt, at mean distances of 2.2 to 3.3 AU from the Sun and with orbital periods of 3 to 6 years. For comparison, the mean distance between the Earth and the Sun is 1 AU, about 150 million kilometers. Not all asteroids lie in this belt, but those that do are said to belong to the main belt.

The Sicilian astronomer Giuseppe Piazzi (1746–1826) discovered the first asteroid accidentally, on 1 January 1801, the first day of the 19th century. He named the tiny object Ceres, in honor of the patron goddess of Sicily. Another asteroid, named Pallas, was located the following year, and the third and fourth, designated Juno and Vesta, were found in 1804 and 1807, respectively.

Ceres is the largest body in the main belt. The mass of this asteroid comprises about one-third the total mass of all the asteroids. It is about 950 kilometers across, and has a water-rich surface. Smaller Pallas and Vesta are about half the size of Ceres and a quarter of the mass. Vesta has experienced significant heating and differentiation, showing signs of a metallic core and a dry surface covered by basaltic lava flows.

We now know that there are very many smaller asteroids. After a gap of nearly a century, astronomical photography enabled them to be found by the hundreds. During a long-exposure photograph made through a telescope tracking the stars, the faster-moving asteroids make short trails while the stars look like dots.

When an asteroid is discovered, it receives a temporary designation, consisting of the year of discovery followed by two letters. The first letter indicates the half-month of the asteroid discovery – the letters "I" and "Z" are not used – to make a total of 24 half-months. The second letter shows the order of discovery within that half-month, but the letter "I" is not used for this second letter. Thus, the asteroid 1998 KY was the 24th (letter Y) asteroid found during the second half of May (letter K) in 1998.

After an asteroid has been observed often enough for an accurate orbit to be established, it receives a number corresponding to the chronological order of reliable orbit determination – but not the order of discovery. The number is often followed by a name provided by the discoverer. As of January 2010, there were 231 665 asteroids with accurate orbits and official numbers, and about 200 000 more have been observed well enough to obtain preliminary orbits.

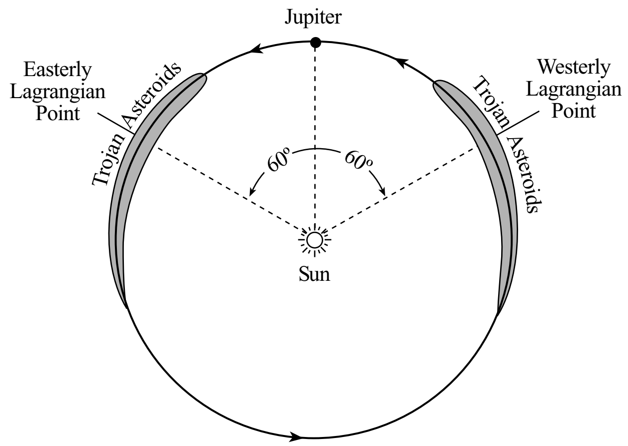


Fig. 12.1 Trojan asteroids Asteroids that share Jupiter's orbit are known as the Trojan asteroids. They are located near the point where the gravitational force of Jupiter and the Sun are equal. The gravitational perturbations of the inner planets produce slight swinging motions, so the Trojan asteroids oscillate within the two shaded regions. Some of the Trojan asteroids may occasionally move close enough to be captured by Jupiter's gravity, thereby accounting for the planet's unusual outer satellites.

With a million asteroids whirling around the Sun, the main belt was once thought to be a hazard to space travel. In reality, however, the volume of space they occupy is so large that any one asteroid is typically several million kilometers away from its nearest neighbor, and the asteroids are so small that it would be hard to hit one even if you tried to do so. Thus, despite their vast numbers, the asteroids leave plenty of room for space flight. The *Pioneer 10* and *11*, *Voyager 1* and *2*, *Ulysses*, *Galileo* and *Cassini* spacecraft have all passed through the asteroid belt unharmed.

And since the asteroid belt is largely empty space, collisions between asteroids are relatively rare over short timescales. It takes tens of millions of years for shattering collisions to occur, so they do happen but it takes a long time.

Trojan asteroids

Not all asteroids are found in the main belt. An especially interesting type is further away from the Sun than the asteroid belt. They move along Jupiter's orbit, keeping pace with the giant planet. The first known one, 588 Achilles, was discovered photographically by the Heidelberg astronomer Max Wolf (1863–1932) in 1906. Hundreds of them are now known, traveling on both sides of Jupiter in two clouds, one preceding the giant planet and one following it (Fig. 12.1). As with Achilles, they are all named after heroes of the Trojan War and they are therefore collectively known as the Trojan asteroids.

The Trojan asteroids are held captive by the gravity of both Jupiter and the Sun. They are found near two of the five Lagrangian points, named after the Italian-born French mathematician Joseph Louis Lagrange (1736–1813) who predicted their existence 134 years before the discovery of 588 Achilles. At these points, the gravitational force of Jupiter is equal to that of the Sun, which is much more massive than Jupiter but also a lot further away from the asteroids. These Lagrangian points lie in the corners of equilateral triangles that have Jupiter and the Sun at the other corners (Fig. 12.1).

The Trojan asteroids do not stay precisely at a Lagrangian point, but instead oscillate around it. They pace back and forth along Jupiter's orbit in paths that take them toward and away from the planet over a cycle lasting hundreds of years.

Once locked into their haven near the Lagrangian points, the Trojan asteroids move at slow speeds along well-defined paths. The Trojans therefore suffer fewer collisions than their counterparts in the main asteroid belt, and their surfaces have probably gone unchanged for billions of years. Thus, the Trojan asteroids may be pristine remnants of the early solar system.

The Trojan asteroids are in a 1 : 1 mean-motion resonance with Jupiter, meaning that they have the same orbital period. Other groups of asteroids are locked into other dynamical resonances with Jupiter. These include the Hildas at the 3 : 2 mean-motion resonances, where the asteroids orbit with two-thirds of Jupiter's orbital period, the Thule group at the 4 : 3 resonances, and the Hungarias group at the 9 : 2 resonances.

Near-Earth asteroids

Although the vast majority of asteroids travel in the main belt lying between the orbits of Mars and Jupiter, there are some notable exceptions that reside within the inner solar system. Known as the near-Earth asteroids, they move inward toward our planet as they travel around the Sun.

One of them could hit our planet someday, with devastating consequences. So astronomers are still actively trying to find all of these interlopers and monitor their motions, with the hope of avoiding the collision. Since these asteroids move closer to the Sun than Mars, they travel at faster speeds than asteroids in the main belt, and thus make longer trails on photographic or computerized images taken to discover asteroids.

There are three populations of near-Earth asteroids, called the Atens, Apollos and Amors (Fig. 12.2). Both the Aten and Apollo asteroids move on eccentric orbits that can cross the Earth's path in space. The Atens are always close to the Sun, never moving out as far as the orbit of Mars. The elongated orbits of the Apollo objects loop in from the main belt to within the Earth's orbit. The Amors travel around the Sun between the orbits of Mars and the Earth, and often cross the orbit of Mars.

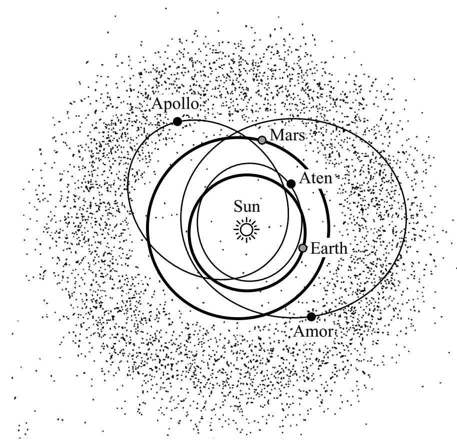


Fig. 12.2 Near-Earth asteroids The paths of three representative near-Earth asteroids, 1221 Amor, 1862 Apollo and 2062 Aten, all come closer to the Sun than most asteroids, located in the main belt beyond the orbit of Mars. Amor crosses the orbit of Mars, and almost reaches the Earth's orbit. Apollo crosses the orbits of Mars, Earth and Venus (not shown). Aten is always fairly close to the Earth's orbit.

Chaotic orbits of asteroids

Why do some asteroids move near the Earth, while most of them stay in the asteroid belt? Gaps of missing asteroids in the main belt provide some clues to these wandering interlopers. These are the Kirkwood gaps (Fig. 12.3), discovered long ago by the American astronomer Daniel Kirkwood (1814–1895).

The locations of these clearings correspond to orbital resonance with Jupiter, in which the orbital periods are exact fractions of the giant planet's period. Any object that orbits the Sun at the 3 : 1 resonance, for example, would have exactly one-third, or 1/3, the orbital period of Jupiter, and it would complete three circuits around the Sun for every one that Jupiter completes. Such an asteroid would revolve around our star at a distance of 2.5 AU with a period of 3.95 years, compared with Jupiter's orbital distance of 5.2 AU and orbital period of 11.86 years. An asteroid that happened to stray into this resonance would come close to Jupiter at almost the same part of the asteroid's orbit at regular 11.86-year intervals and the accumulated gravitational interaction with Jupiter could dislodge the asteroid from its orbit.

So the net effect of the resonance is to clear asteroids out of the resonant locations, and some of them might eventually be brought into Earth-crossing orbits. Asteroids

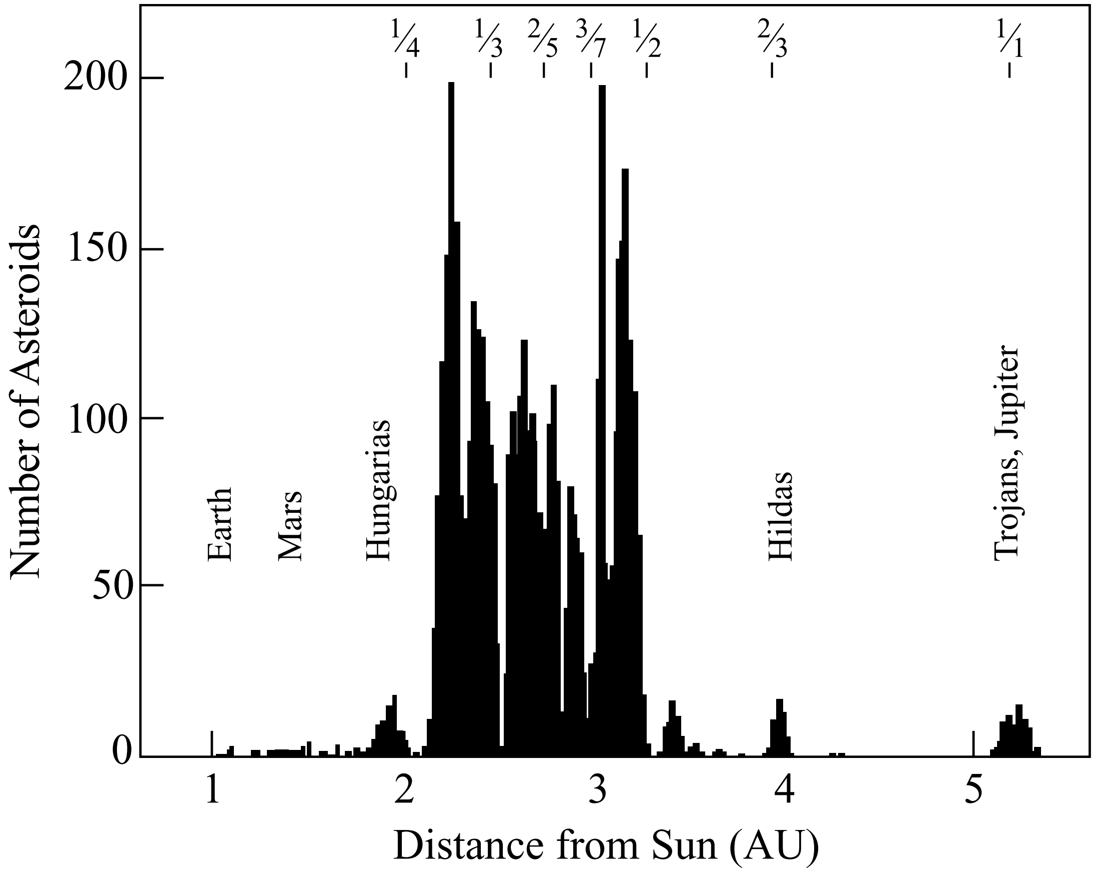


Fig. 12.3 Kirkwood gaps in asteroid orbits Most of the asteroids are found in the asteroid belt that lies between 2.2 and 3.3 AU from the Sun, but there are very few asteroids at certain distances from the Sun. These Kirkwood gaps have orbital periods of 1/4, 1/3, 2/5, 3/7 and 1/2 of Jupiter's orbital period. These fractions are placed above the relevant gap in the figure. The fraction 1/4 indicates, for example, that an asteroid at that distance makes four revolutions around the Sun for each one revolution completed by Jupiter. Asteroids are tossed out of such resonant orbits by Jupiter's repeated gravitational perturbations. In addition, there are several peaks corresponding to groups of asteroids with nearly the same orbital distance, such as the Trojan asteroids that have orbits that are identical in size to the orbit of Jupiter.

that are not in a resonance are affected by Jupiter at completely random time intervals and places along their orbit, so the giant planet's gravitational perturbations tend to cancel each other over long times. It is somewhat analogous to repeated pushes on a swing. If the pushes occur at the same point in each swing, they can amplify the change, but haphazard pushes would produce little net effect.

A satisfactory explanation of the Kirkwood gaps was not achieved until the 1980s when powerful computers were used to study Jupiter's influence on the motion of asteroids. The computer simulations showed that Jupiter induces a chaotic zone in the vicinity of an orbital resonance, and that an asteroid that moves into the resonant orbit will eventually be tossed out of it. An asteroid in the chaotic zone can spend tens of thousands of years in a well-behaved, near-circular orbit. But that ordered placid behavior could be unexpectedly interrupted after 100 000 years or so, when the orbit is suddenly stretched and elongated in a chaotic way (Fig. 12.4).

The increasingly elongated orbit can become so elliptical that it crosses the orbit of Mars or the Earth, and gravitational interaction with these planets can fling the asteroid into a totally different trajectory that removes it from the asteroid belt. Thus, orbits that initially fall within the chaotic zone around a resonance with Jupiter would be gradually cleaned out, creating a gap in the asteroid belt. Asteroids that have been thus redirected could abruptly end their voyage through space in a collision with Earth.

12.2 Origin of the asteroids

Former worlds

In the past, there have been two extreme theories for the origin of asteroids. According to the first, the asteroids represent the fragments of a former planet that has been torn apart. The second theory proposes that the asteroids are the

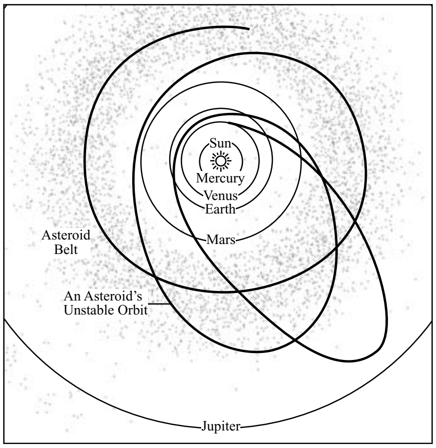


Fig. 12.4 Chaotic asteroid orbit Asteroid orbits can become chaotic under the gravitational influence of nearby massive Jupiter. Asteroids at certain locations in the main belt follow a trajectory that becomes increasingly off-center over thousands of orbits. They may eventually become interlopers with orbits that cross the Earth's orbit. Some of these near-Earth asteroids may eventually collide with our planet.

pieces of a planet that never formed. Today, astronomers favor a theory that lies between the extremes.

It is now known that the combined mass of all asteroids is far too small to make up a major planet. If all the known asteroids were brought together, they would create a body less than 5 percent the mass of the Earth's Moon. So the first extreme must be discarded. On the other hand, the second extreme can also be excluded because there is strong evidence that many of the asteroids were once collected into a relatively small number of slightly larger parent bodies.

Remnants of a planet that never formed

Why did the asteroids fail to coalesce into a single planet? It is likely that gravitational forces from the rapidly forming and massive Jupiter took charge of its neighborhood, stirring it up and keeping the original asteroids from growing too large. Numerous asteroids in resonant orbits with youthful Jupiter probably permeated the region of the main asteroid belt. Chaotic zones in the vicinity of these resonances would have pumped up the eccentricities of initially circular orbits, flinging the resident bodies into elongated and inclined orbits, accelerating them to high velocities, and causing them to crash into each other. The colliding objects would be moving too fast to stick to each other. Instead, they would break apart into fragments.

Collisions therefore pulverized the early asteroids, grinding them down to the numerous smaller asteroids we see today. Almost every asteroid we see today must be a fragment of a larger original body.

In contrast, a swarm of small, solid bodies in the inner solar system, with orbits closer to the Sun than the asteroid belt, were located far from Jupiter's gravitational influence. These so-called planetesimals therefore remained in nearly circular orbits, moving at slow velocity around the Sun. This permitted the neighboring planetesimals to merge gently together and form larger ones. They eventually coalesced into the four terrestrial planets – Mercury, Venus, Earth and Mars.

If it wasn't for Jupiter's chaotic interference, a similar terrestrial planet might have formed in the asteroid belt. Gravitational interaction with the giant planet removed objects from this region, throwing them into collisions with the other planets or out of the solar system. These events cleared the asteroid zone of 99.9 percent of its original mass. The missing material could have coalesced to form an Earth-sized planet. What remain today are the relic building blocks of a planet that failed to form.

A lifetime of catastrophic collisions

Billions of years ago, before Jupiter began to disrupt nearby objects, a few large bodies probably inhabited the asteroid belt. The largest of these would-be planets accumulated enough internal heat to differentiate, their dense material sinking to form iron cores and leaving rocky residues in their outer layers. Volcanic flow may have covered the surface of some of them.

Most of the initial asteroids never grew large enough to form cores, so they preserved matter typical of the region in which they formed. The majority of asteroids, both large and small, are undifferentiated asteroids.

Encounters among the earliest asteroids became increasingly violent as Jupiter stretched and twisted their orbits into eccentric and inclined orientations. These orbits crossed each other, sometimes resulting in violent collisions. Instead of continuing to grow, the largest asteroids were chiseled and blasted apart by mutual collisions.

So the original asteroids never grew larger than about 1000 kilometers across, and they never accumulated into a major planet. The pulverized remnants of these former worlds became the present asteroids, often orbiting the Sun in families with common orbital characteristics and spectral properties.

The Japanese astronomer Kiyotsugu Hirayama (1874–1943) discovered asteroid families. He noticed that groups of main-belt asteroids share very similar orbits (Fig. 12.5), suggesting that they are the broken fragments of larger objects. Hirayama called each group a family because he believed the members shared a common origin as the

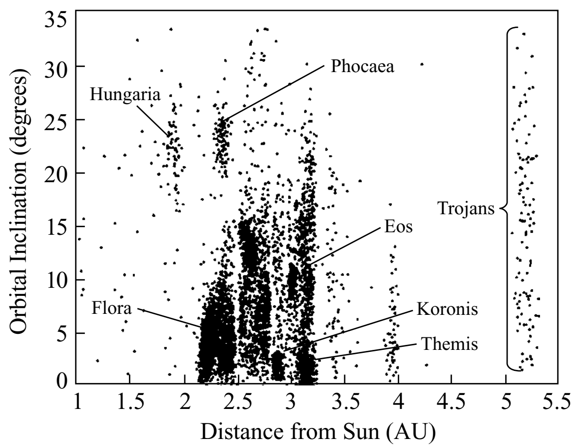


Fig. 12.5 Asteroid families The orbital parameters of many asteroids are very similar, as shown in this diagram of orbital inclination (*vertical axis*) plotted as a function of orbital distance from the Sun (*horizontal axis*). Three families with common orbital elements are the Koronis, Eos and Themis families. The Flora family is sometimes subdivided into several separate ones. The Hungaria and Phocaea group of asteroids at high inclinations are separated from the main belt by secular resonance with Jupiter, that clears asteroids out of certain locations, and they are not true families. The Kirkwood gaps, also cleared by resonance, are noticeable by vertical white spaces; the one at 2.5 AU corresponds to an orbital period of 4 years and the 3 : 1 resonance. Another sharp break is present at 3.3 AU, corresponding to a period of 6 years and the 2 : 1 resonance. [Adapted from Charles T. Kowal's *Asteroids* (New York: John Wiley, 1996).]

children of a bigger parent body. He also named a number of families after their largest member asteroids, such as the Eos, Koronis and Themis families.

Within each of these families, the orbits are so similar that the members must have originated from a single object. Hundreds, and perhaps thousands, of small asteroids making up each family are probably the debris of a collision that disrupted the once-larger parent body. These parents may have been several hundred kilometers in diameter.

And not only are the orbits similar within each family, the colors and surface compositions are also often alike. These similarities imply that the families are real physical groupings. The Koronis family is, for example, composed of bright silicate asteroids and the Themis family of dark carbonaceous ones. Some family members are spectrally diverse, perhaps sampling different parts of the interiors of their parental precursors.

The asteroid belt has been dominated by collisions for billions of years. Major impacts between asteroids larger than one kilometer across occur every 10 to 100 million years. So each one of them must have suffered roughly a hundred devastating collisions over the past 4 billion years.

Over time, the asteroids crashed into one another and modified their shapes. When a small asteroid hit a larger one, it gouged a crater out of the surface of the larger one. If the two colliding objects were of comparable size, they could have broken each other apart into smaller fragments. As a result, some present-day asteroids are the metal-rich cores of larger former parents, stripped of their rocky mantles by the ongoing collisions. Others are the shattered remains of bodies that have remained homogeneous. Such destructive encounters occur even now, dominating the evolution, shapes and sizes of the asteroids.

12.3 Viewing asteroids from a distance

The size of asteroids

Due to its small size, an asteroid remains an unresolved point of light in even the best telescopes on Earth, just like a faint star. This explains the name *asteroid*, which comes from a Greek word that means "starlike". Although the name describes the visual appearance of these objects in a telescope, it is totally inappropriate to their physical nature. Using our instruments on Earth, we can determine the sizes of asteroids, and they are much smaller than either a star or a major planet.

One method of measuring the size of an asteroid is to watch it pass in front of a star. During such a stellar occultation, the asteroid casts a shadow on the Earth, and the width of this shadow is equal to the asteroid's diameter. As the Earth rotates and the asteroid moves across the sky, the asteroid's shadow is swept by an observer on Earth, and its width can be determined by measuring the length of time the star is invisible. Observations of stellar occultations have established accurate values for the sizes of the largest asteroids, such as 1 Ceres and 2 Pallas (Table 12.1).

Even the biggest asteroids are smaller and less massive than the Moon. Ceres is by far the largest asteroid, having a radius of about 475 kilometers and a mass of 9.43×10^{20} kilograms. That is about a third the radius of the Earth's Moon and only 0.016 the mass of the Moon. The brightest asteroid, Vesta, is about half the size of Ceres and about the same size as Pallas.

There are many more small asteroids than big ones. About 1000 asteroids are larger than 15 kilometers in radius. Surveys of the faintest asteroids suggest that there are about half a million asteroids in the main belt larger than 1.6 kilometers. Yet, despite their vast numbers, the total mass obtained by adding up the contributions of all asteroids, of all sizes, is far less than the mass of any major

Table 12.1 Principal characteristics of the three largest asteroids[a]

Asteroid	Type	Radius (km)	Mass[b] (10^{21} kg)	Bulk density (kg m^{-3})	Rotation period (hours)	Semi-major axis (AU)
1 Ceres	C	475	0.943	2120 ± 40	9.1	2.77
2 Pallas	B	266[c]	0.214	2710 ± 110	7.8	2.77
4 Vesta	V	265	0.267	3440 ± 120	5.3	2.36

[a] The radius is in units of kilometers (km), the mass is in units of kilograms (kg), and the bulk mass density is in units of kilograms per cubic meter (kg m^{-3}).
[b] The mass determinations are from the work of E. M. Standish at the Jet Propulsion Laboratory where he uses observations of Mars to solve for the masses of the largest asteroids.
[c] Pallas is not quite spherical. A tri-axial ellipsoid fit to occultation observations has diameters of 559 km, 525 km, and 532 km with a mean diameter of 538 ± 12 km.

planet. The entire asteroid belt is only 0.05 the mass of our Moon and just 0.0006 the mass of the Earth.

Another way of estimating an asteroid's size is to measure its apparent brightness. Bigger asteroids reflect more sunlight and are therefore brighter. The amount of sunlight reflected from an asteroid nevertheless depends on both its size and reflectivity, or albedo: this is the ratio of light reflected by a surface to the incident light, and therefore is a measure of the efficiency of the reflection process. As an example, 1 Ceres is the biggest asteroid, but 4 Vesta can appear brighter because it has a higher albedo. Vesta reflects 42 percent of the sunlight that strikes it, compared with Ceres' reflectivity of only about 9 percent.

Both the size and the albedo of an asteroid can be determined by observing it in visible and infrared light. When sunlight strikes an asteroid, some of the light is reflected, but most of it is absorbed. The absorbed radiation makes the surface heat up and emit infrared light. By measuring the infrared radiation from the asteroid and comparing it to the amount of reflected visible light, we can determine its albedo. And by knowing the asteroid's apparent brightness and distance as well, we can infer its size.

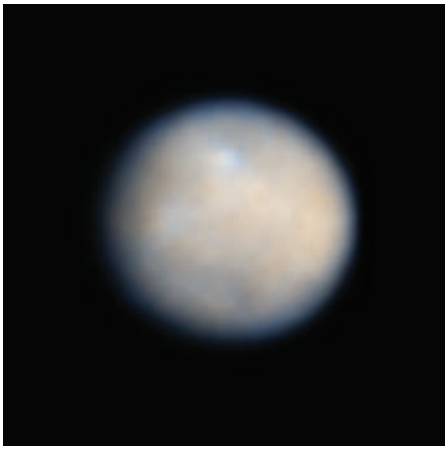


Fig. 12.6 Dwarf planet-asteroid 1 Ceres A *Hubble Space Telescope* image of 1 Ceres, the largest object in the asteroid belt and the first object discovered there. The round shape led to its classification in 2006 as a dwarf planet, and suggested that its interior is layered, with a rocky core, icy mantle and thin, dusty outer crust. Ceres is approximately 950 kilometers across. The differences between lighter and darker regions could be impact features or simply due to a variation in surface material. (Courtesy of NASA/ESA/J. Parker, SRI, P. Thomas, Cornell U., L. McFadden, U. Maryland, and M. Mutchler and Z. Levay, STScI.)

Shape and form of asteroids

The rotation of a non-spherical asteroid can bring various surfaces into view, reflecting different amounts of sunlight. This produces a brightness variation, with periods of hours, that can be used to infer an asteroid's irregular shape. These rotational modulations of the light reflected from the largest asteroid, 1 Ceres, are very small, indicating that this asteroid is practically spherical in shape.

The round shape of 1 Ceres has been confirmed from images taken with the *Hubble Space Telescope* (Fig. 12.6). Ceres is massive enough that its gravity has crushed the object into this spherical form. The International Astronomical Union has therefore added a new designation to Ceres, classifying it as a dwarf planet since it is small, orbits the Sun, and has sufficient mass to attain a nearly round shape.

Asteroids are also known as minor planets, but they usually have irregular shapes and differ noticeably from the familiar major planets. Because of their small size, asteroids have very little mass or gravity. Unlike the larger planets, asteroids are not big enough to hold onto an atmosphere or to retain a spherical shape. The exception is Ceres.

Heat from formation and the decay of radioactive elements most likely elevated the interior temperature of Ceres to the point of differentiation, in which heavier, denser material sank to the center while lighter materials rose to the surface. Scientists therefore think that Ceres has a rocky inner core surrounded by a thick water-ice mantle and a thin, dusty outer crust.

Most, if not all, of the smaller, less massive asteroids have irregular shapes, and retain their asteroid classification. Their self-gravity is not sufficient to overcome internal rigid body forces. Their irregular shapes cause pronounced rotational brightness variations.

The majority of asteroids are not large enough to have ever had a molten core, and lava has never flowed out from the inside of most of them. So, except for ongoing collisions, most asteroids have been unaltered since the formation of the solar system about 4.6 billion years ago. In contrast, the surfaces of large planetary bodies have been destroyed by eons of planetary evolution, including immense volcanoes and extensive lava flows across their surfaces.

The form and amplitude of the periodic light variations have demonstrated that most asteroids are at least slightly elongated, chipped and pummeled into irregular shapes. They are too small to retain a spherical shape during their lifetime of disruptive collisions.

The stretched-out irregular shapes of some asteroids have also been determined from radar observations of near-Earth asteroids that travel close enough for scientists to detect the echoes of radio waves bounced off them. During their close approach, these asteroids speed by the Earth at distances of several hundred thousand kilometers, permitting brief high-resolution radar images before they move on and fade from view.

Radar is an acronym for Radio Detection and Ranging. In this technique, a pulsed radio signal is sent to the asteroid, and the properties of the return "echo" are compared to those of the transmitted signal. This yields information about the composition, surface texture, shape and rotation of the asteroid. Smooth surfaces reflect more radio energy than rough ones, and the faster an asteroid rotates the broader the range of returned wavelengths, which have been Doppler-shifted by the rotation to longer and shorter waves.

The radar data indicate that the overall shape of some asteroids is dominated by two irregular, lumpy components that touch each other, something like a dumbbell. Such an asteroid is a double object, that is, two bodies in contact. The two pieces probably merged after a past catastrophic collision of a larger body; they may have been thrown apart and subsequently came together under their mutual gravity. Or they might be two former asteroids that joined in a gentle encounter. Radar images have also revealed at least one triple asteroid, a central body about 700 meters across orbited by two smaller moons.

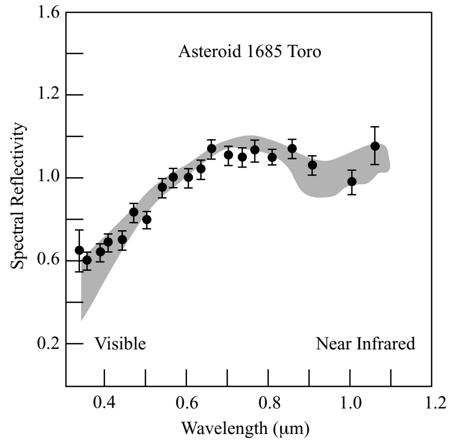


Fig. 12.7 Finding the composition of asteroids A prominent silicate absorption feature is present in the spectrum of sunlight reflected from the S-type asteroid 1685 Toro (*dark dots with error bars*). Asteroids like Toro may be the source of the stony meteorites recovered on Earth. The shaded spectrum is the reflection spectrum of a stony-chondrite meteorite. The wavelength is in units of micrometers (μm), or 10^{-6} meters.

An asteroid's color

Because asteroids display no visible disk when observed from Earth, their physical characteristics must be inferred from the intensities and spectral properties of their reflected sunlight. By comparing an asteroid's reflected light, wavelength by wavelength, with that of the incident sunlight, it is possible to deduce its surface composition. Astronomers divide the amount of incident sunlight at each wavelength by the amount of reflected sunlight at that wavelength, and the ratio tells them how much light of each color is reflected compared to any other color. Such spectral measurements have revealed the physical diversity of the asteroids, and shown that their compositional differences tend to depend on distance from the Sun.

The bulk of main-belt asteroids can be divided into two broad spectral categories, known as the S, for silicate, and C, for carbonaceous, types. The bright S-types have a reddish color and exhibit spectral dips identified with absorption by silicate minerals (Fig. 12.7). They prevail in the inner part of the asteroid belt, orbiting closer to the Sun than the belt mid-point (Fig. 12.8).

In contrast, the C-type asteroids are darker, bluer and richer in carbon, with relatively flat and featureless spectra

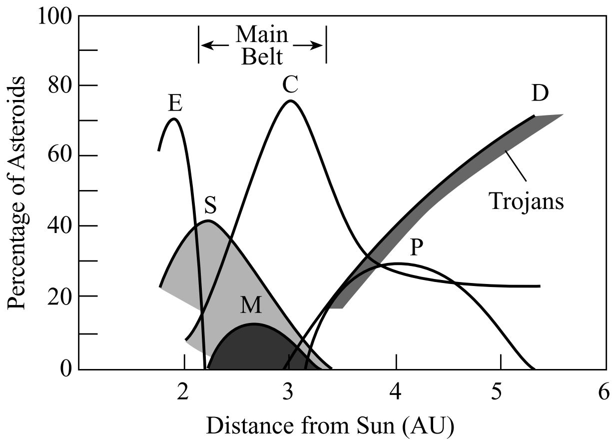


Fig. 12.8 Asteroid distribution of spectral type with distance
The color, or surface composition, of the asteroids is correlated with distance from the Sun. In order of increasing distance, there are the white E asteroids, the reddish S or silicate ones, the black C or carbonaceous ones, and the unusually red D asteroids. This systematic change has been attributed to a progressive decrease in temperature with distance from the Sun at the time the asteroids formed. The rare metallic M asteroids are probably the cores of former, larger parent bodies. The P asteroids have spectra that are indistinguishable from the M ones, but the P types are darker.

Focus 12.1 Mining asteroids

The asteroids are rich storehouses of valuable materials such as iron, nickel and water. The utilization of their minerals could overcome growth limits imposed by dwindling natural resources on Earth.

Prospecting spacecraft might be sent to an asteroid in search of a cosmic El Dorado. Because of its low gravity, valuable metals could be easily removed, and the extracted material might be shipped back by spacecraft from an asteroid that traveled near to the Earth. Imaginative engineers speculate that a more distant asteroid might be brought closer to Earth using a "mass-driver", a device that would chew off pieces and fling them into space, propelling the asteroid like a rocket.

Metals are relatively scarce in the Earth's outer layers because most of them sank inside to the central core where they will remain forever sequestered, buried under our planet's crust and mantle. In contrast, some metallic asteroids may be pure core. The outer mantle and crust of their former parents were probably chipped away by eons of cosmic collisions. So the metals on an asteroid could be relatively abundant and easy to extract. Moreover, the residual gas and dust would be swept from the solar system by the solar wind, thereby avoiding the problems of industrial pollution when recovering valuable metals on Earth.

at visual wavelengths. The C-type asteroids far outnumber all types, possibly composing three-fourths of the main belt. The C-type asteroid 1 Ceres is a representative example; it has a smooth visible spectrum with infrared absorption features attributed to clay-like minerals that include water as a part of their molecular structure. The iron-rich clays may have formed on what was once a wet surface.

The S-type asteroids probably account for up to 15 percent of all asteroids. Some of the less common M-type asteroids reflect sunlight in a way that suggests that their surfaces are composed of nickel and iron, hence the designation M for metallic for at least some of them. These objects could be the metal cores of larger parent bodies, stripped down to their cores by collisions. The M-types are most common in the middle of the asteroid belt. Future space entrepreneurs may want to mine them for valuable metals (Focus 12.1).

There is an intriguing connection between the composition of asteroids and their distance from the Sun. The innermost asteroids, with orbits closest to the Sun, are rocky, siliceous and dry, while the outer ones are carbonaceous with water-rich, clay-like minerals.

There is a related, progressive decrease in an asteroid's reflecting power with increasing distance from the Sun. The brightest asteroids that reflect the most sunlight tend to lie near the inner edge of the main belt, closest to the Sun, while the most distant asteroids are, on the average, the darker ones with the lower reflecting power.

These differences in the composition and reflecting power of asteroids are probably related to conditions in the primeval solar nebula – the interstellar cloud of gas and dust from which the solar system originated. They may be a consequence of a decrease in temperature with increasing distance from the young Sun when the asteroids were formed.

The middle of the main asteroid belt marks the boundary dividing the cold, outer, water-condensing regions and the hot, inner, water-vapor parts of the primeval nebula. Dark material, rich in carbon and water, could condense in the colder regions farther from the Sun, but not in the hot regions near to our star where the early water was probably vaporized. On the other hand, the bright rocky material was less volatile and could remain within the hotter regions closer to the Sun. In this way, the temperature of the solar nebula may have led to the pattern of materials and colors now seen in the asteroids.

Astronomers have identified a plethora of other, less-common classes of asteroids, based upon the shape and slope of their reflectance spectra. In addition to the most common S-, C- and M-types, there are at least 11 other classes denoted by different letters, such as the red, possibly organic D-types found in the outer belt, and the white E-types that are closer to the Sun. In addition, rare individual asteroids exhibit unique well-identified absorption features in their spectra. The brightest and third largest of the asteroids, 4 Vesta, for example, shows the distinct absorption signature of volcanic basalt. Its parent body spawned a little family of smaller V-type asteroids with similar spectra and basaltic composition.

The spin of an asteroid

Asteroids do not shine like a steady beacon with constant brightness. They instead reflect a varying amount of sunlight toward the Earth. The observed brightness variation, also known as a light curve, is periodic, often with two maxima and two minima. The overall repetition is due to rotation, while the double pattern of variability results from alternating side views of an asteroid's elongated shape. When we see the biggest side of an asteroid, with the greatest area, the asteroid is brightest, while the smaller area reflects less sunlight and the asteroid is dimmer.

Almost all asteroids spin about a single axis. The period of rotation is inferred from the amount of time that it takes for the complete pattern of brightness variation to repeat itself. The rotation periods are usually between 2.4 and 24 hours (Fig. 12.9), although a few of them rotate with longer periods, such as 253 Mathilde with a rotation period of 17.4 days, and some have periods of only a few minutes. Frequent oblique collisions can increase the rate of rotation or decrease it, depending on whether the collision is in the direction of rotation or opposite to it.

Some asteroids are probably rotating as fast as they can. If an asteroid is not solid, and is thus bound only by its own gravity, it can only spin at a certain maximum rate before material is whirled off it. Asteroids larger than 200 meters seem to have reached this limit, for most of them do not rotate faster than once every 2.4 hours (Fig. 12.9), suggesting that there is nothing stronger than gravity holding them together. If it lacks the material strength of a solid, an asteroid with a faster rotation rate and a shorter period will throw material off its surface and fly apart.

Asteroids smaller than 200 meters in diameter can rotate at faster rates, some turning once every few minutes. Their rotation is too rapid for these asteroids to consist of multiple components bound together by mutual gravitation. They must instead be rock solid. Some small asteroids rotate so swiftly that their day ends almost as soon as it begins. Asteroid 2008 HJ has one of the fastest known rotation periods at just 42.7 seconds; it is thought to be a compact, massive object about 20 meters across.

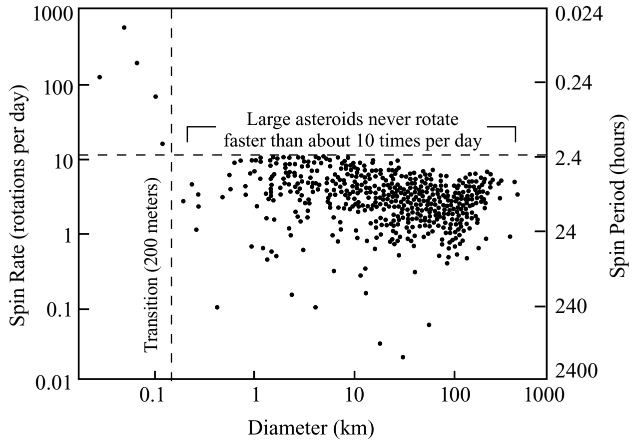


Fig. 12.9 How fast do asteroids spin? Two groups of asteroids are indicated on this plot of their rotation rates (*vertical axis*) versus size (*horizontal axis*). There are the large asteroids that rotate at a wide variety of speeds, except at the fastest rates, and the small asteroids that spin the fastest. No known asteroid larger than 200 meters across rotates faster than once every 2.4 hours, perhaps because these asteroids are piles of rubble that fly apart if spun too fast. Smaller asteroids, which can turn once every few minutes, must be solid rocks.

The direction of an asteroid's rotation axis can be determined by noting the way the brightness variation changes as the asteroid moves about the Sun. A highly elongated asteroid will reveal little or no light variation if its spin axis points directly toward us. As the asteroid's orbital motion carries it away from this alignment, the amount of light reflected toward Earth will increase, reaching a maximum when its broadest equatorial surface is turned toward Earth. Studies of this effect suggest that the rotation axes of the asteroids are haphazardly oriented, pointing in all directions in space, contrary to the orderly rotations of most of the major planets whose rotation axes point more or less north–south. Frequent collisions may have been responsible for the diverse orientations of asteroid rotation axes.

12.4 Spacecraft view asteroids close up

Flyby missions of asteroids

A nearby high-resolution view was required to examine the surface details of asteroids, and this has been accomplished by several passing spacecraft on their way to accomplish a different goal. The first close encounters were provided when the *Galileo* spacecraft flew by two asteroids, 951 Gaspra and 243 Ida, on its way to Jupiter (also see Section 2.1). The images, taken on 29 October 1991 and

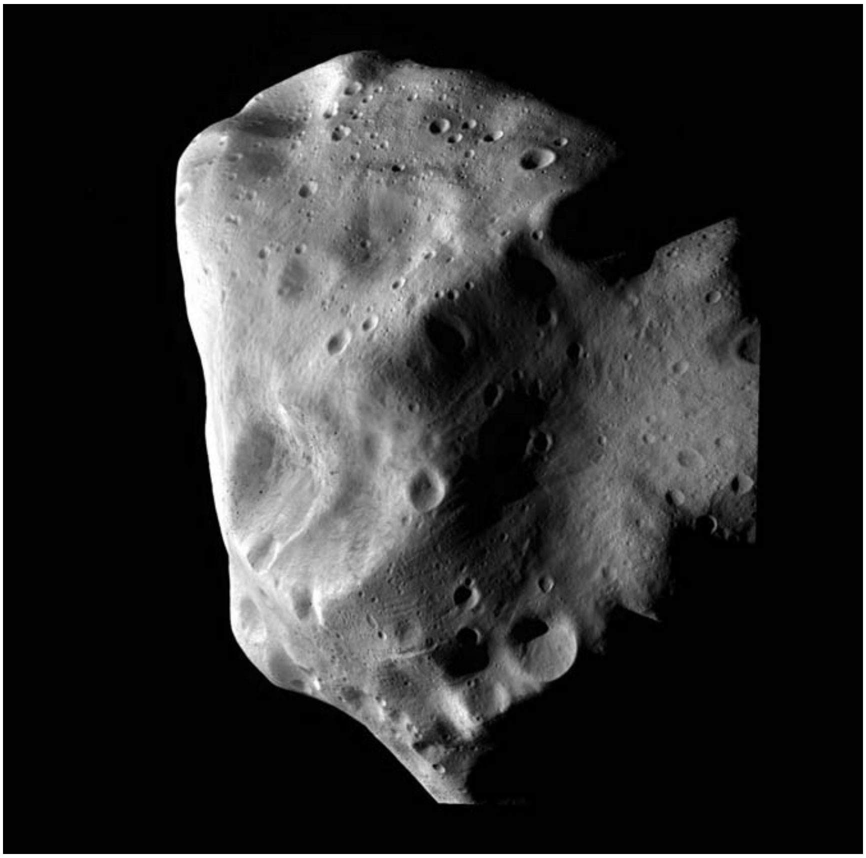


Fig. 12.10 Asteroid 21 Lutetia An instrument aboard the *Rosetta* spacecraft captured his image of the main-belt asteroid 21 Lutetia on 10 July 2010. *Rosetta* is expected to orbit and then land on comet Churyumov–Gerasimenko in 2014. (Courtesy of ESA/*Rosetta* team.)

28 August 1993, respectively, showed irregular shapes peppered by an abundance of small craters, suggesting that asteroids are the battered remains of larger bodies subject to numerous collisions with objects of comparable or much smaller size.

Somewhat blurry images were obtained during the *Deep Space 1* flyby of asteroid 9969 Braille, on 28 July 1999 while testing new technologies, and when the *Stardust* spacecraft passed asteroid 5535 Annefrank on 2 November 2002, en route to its primary target, comet Wild 2. As expected, both asteroids are irregular, elongated bodies, several kilometers across on their longest sides.

On 5 September 2008 the *Rosetta* spacecraft flew by the main-belt asteroid 2867 Steins at a distance of about 800 kilometers. These were the first close-up views of a rare E-type asteroid, classified from its high albedo and infrared spectrum. The *Rosetta* images reveal a conical, oblate diamond-shaped body, with dimensions between 4.5 and 6.7 kilometers, covered with shallow craters. *Rosetta* flew by the large main-belt asteroid 21 Lutetia on 10 July 2010, revealing an irregularly shaped, cratered object about 100 kilometers across (Fig. 12.10). Data obtained during the encounter will help determine if Lutetia is a primitive undifferentiated C-type asteroid or a metallic M-type asteroid from the core of a larger differentiated asteroid. Lutetia is the Latin name for Paris. The *Rosetta* spacecraft is expected to orbit its primary scientific target, comet Churyumov-Gerasimenko, beginning in 2014.

A sideways glimpse of 253 Mathilde, a dark carbon-rich C-type asteroid, was acquired on 27 July 1997 from the *NEAR Shoemaker* spacecraft on its way to orbit asteroid 433 Eros. Like the other asteroids seen close up, Mathilde has survived blow after blow of cosmic impacts (Fig. 12.11). Its surface is covered with the crater scars of past collisions that have disfigured the asteroid's shape, like the battered and scarred face of a professional boxer who has just lost a fight. Huge pieces have been removed from Mathilde, leaving four enormous craters tens of kilometers across.

Over billions of years, asteroids on intersecting orbits have collided with enough force to shatter and break 253 Mathilde into pieces. Instead of dispersing, the fragments have apparently re-accumulated into a rubble pile. The mass of the asteroid was determined by radio tracking of small perturbations in the trajectory of *NEAR Shoemaker* during its close flyby. A similar technique established an accurate mass for 433 Eros and 25143 Itokawa when spacecraft circled them. These mass determinations were combined with measurements of the asteroid volume to provide their bulk mass density (Table 12.2). *Galileo* flew by too fast and too far away to be affected noticeably by the asteroids it encountered, so masses could not be determined in this way. But a rough estimate for the mass of

Table 12.2 Physical properties of asteroids visited by spacecraft[a]

Asteroid	Type	Overall dimensions (km)	Mass (10^{15} kg)	Bulk mass density (kg m^{-3})	Rotation period
21 Lutetia	M	$132 \times 101 \times 26$	1700	3400	8.17 hours
243 Ida	S	$60 \times 25 \times 19$	42 ± 6	2600 ± 500	4.63 hours
253 Mathilde	C	$66 \times 48 \times 46$	103.3 ± 4.4	1300 ± 300	17.4 days
433 Eros	S	$31 \times 13 \times 13$	6.687 ± 0.003	2670 ± 30	5.27 hours
951 Gaspra	S	$18 \times 11 \times 9$	–	–	7.04 hours
25143 Itokawa	S	$0.535 \times 0.294 \times 0.209$	$3.58 \times 10^{-5} \pm 0.018 \times 10^{-5}$	1950 ± 140	12 hours

[a] The overall dimensions in kilometers (km) are the diameters of a triaxial ellipsoid fit. The mass is in kilograms and the bulk mass density is in units of kilograms per cubic meter (kg m^{-3}).

Fig. 12.11 Asteroid 253 Mathilde An image mosaic of the C-type asteroid 253 Mathilde acquired by the *NEAR Shoemaker* spacecraft on 27 June 1997. The part of the asteroid shown is about 59 kilometers by 47 kilometers across. The angular shape of the upper left edge is a large crater viewed edge on. Mathilde's low mass density indicates that it is a porous rubble pile. (Courtesy of NASA/JHUAPL.)

243 Ida was inferred from the motion of its tiny moon Dactyl.

Mathilde has an exceptionally low bulk mass density of just 1300 kilograms per cubic meter, only slightly higher than water at 1000 in the same units, and about half the average density of the Earth's crust. The asteroid must therefore contain as much empty space as rock in its interior. It is indeed a loose pile of rubble, broken apart and stuck back together again, so pervasively fragmented that no solid bedrock is left.

The porous interior of Mathilde might explain the mysterious absence of visible rims or ejected deposits around its enormous craters. The unusually large craters may have been formed by compression of the surface during impact, like the dents in a beanbag, rather than by excavation of the material. Mathilde might even have swallowed up the colliding objects, like a bullet shot into a sandbag.

We will next turn to the results of two asteroid-orbiting spacecraft, *NEAR Shoemaker* for 433 Eros and *Hayabusa* for 25143 Itokawa. Eros is solid inside, held together by its own material strength. In contrast, Itokawa seems to be a rubble pile, the reassembled debris of previous impacts, consisting of smaller pieces loosely held together by their mutual gravitation.

NEAR embraces asteroid 433 Eros, a solid rock

On Valentine's Day, 14 February 2000, the *Near Earth Asteroid Rendezvous* (*NEAR*) spacecraft became the first to orbit an asteroid, arriving at 433 Eros, a bright silicate S-type asteroid, after a four-year journey from Earth. It studied the shape and form of the asteroid's surface in high resolution, and determined its composition, mass and bulk density before landing on the surface nearly one year after beginning orbit.

The *NEAR* craft was the first in NASA's Discovery Program of no-frills, scientifically focused, low-cost missions, designed to do quality science in a "faster, better, cheaper" mode. The mission took just 26 months from start to launch, at a bargain total cost of $224 million. The car-sized vehicle circled Eros for a year, landing on the asteroid on 12 February 2001, another historic first. In the meantime, NASA renamed the spacecraft *NEAR Shoemaker* in honor of the astronomer–geologist Eugene M. Shoemaker (1928–1997), a pioneering expert on asteroid and comet impacts.

Discovered in 1898, asteroid 433 Eros was one of the first asteroids found to cross the orbit of Mars and come near the Earth's path in space. When *NEAR Shoemaker* arrived at the asteroid, it was just 1.7 AU from Earth, or less than twice the average distance between Earth and the Sun. Eros is one of the largest of the near-Earth asteroids, about twice the size of Paris and with a mass thousands of times greater than other asteroids that approach the

Fig. 12.12 Asteroid 433 Eros, a solid rock This global view of the S-type asteroid 433 Eros was obtained by the *NEAR Shoemaker* spacecraft on 29 February 2000 from a distance of 200 kilometers. This perspective highlights the major features of the asteroid's northern hemisphere. The asteroid's largest crater (*top*) measures 5.5 kilometers wide and sits opposite from an even larger 10-kilometer, saddle-shaped depression (*bottom*). Studies of the spacecraft's orbit around the asteroid indicate that 433 Eros is a solid rock. (Courtesy of NASA/JHUAPL.)

Earth. So Eros is big and relatively nearby, two important reasons for a visit. It is also named after the Greek god of love – most appropriate for an encounter on Valentine's Day.

Radio-tracking of the orbiting spacecraft was used to determine the mass of Eros, which weighed in at 6.687×10^{15}, kilograms, about one-billionth the mass of Earth. That means that most adults would weigh less than a few ounces if standing on Eros, about as heavy as a bag of airline peanuts on Earth. And on Eros you could jump thousands of meters high, never to return. The gravity is so slight that *NEAR Shoemaker* had to keep its speed down to about 5 kilometers per hour to stay in orbit, moving about as fast as a casual bicyclist. If it moved at a faster speed, the spacecraft would escape the asteroid's feeble gravity and move into interplanetary space.

Although previous spacecraft passed close to a few asteroids, none had orbited one. In contrast to these previous brief flybys, *NEAR Shoemaker*'s cameras scrutinized 433 Eros for a solid year, sending back 160 000 images of the asteroid and recording its diverse surface from all angles and distances.

Eros is a warped and misshapen world, with heavily cratered expanses abutting relatively smooth areas (Fig. 12.12). The asteroid's biggest crater measures 5.5 kilometers across, and most of the surface is peppered with smaller craters (Fig. 12.13). The *NEAR* scientists spotted thousands of them.

The craters on Eros have been given names of famous lovers from history and fiction, taken from different cultures. They include Bovary from Gustav Flaubert's (1821–1880) novel *Madame Bovary*, Don Quixote and his Dulcinea from Miguel de Cervantes' (1547–1616) novel *Don Quixote de la Mancha*, Lolita from Vladimir Nabokov's (1899–1977) novel of that name, Don Juan, the legendary Spanish nobleman known for his seduction of women, and both Eurydice and her lover Orpheus, from Greek mythology.

The surface of Eros is saturated with craters. There are as many craters on its surface as there can be, so continued cratering would not change the overall appearance of its surface. It has been completely formed and shaped by the unrelenting bombardment of impacts over billions of years.

There are nevertheless an unusually low number of small craters pockmarking Eros. Scientists think that shaking and vibrations resulting from impacts have obliterated the small craters, filling those smaller than 100 meters across with rock and dust. They call this erasure *seismic*

Fig. 12.13 Craters on asteroid 433 Eros The many craters on the surface of the asteroid 433 Eros are attributed to eons of collisions with other asteroids. Large boulders, perhaps broken off Eros during these impacts, are perched on one of the crater's edge. The largest boulder, on the horizon in the center of the picture, is about 40 meters long. The two overlapping craters shown here were probably formed many millions of years apart. This picture, taken on 7 July 2000 from the *NEAR Shoemaker* spacecraft, is 1.8 kilometers wide. (Courtesy of NASA/JHUAPL.)

shaping, in analogy with terrestrial earthquakes that shake the ground with seismic waves, but from a totally different cause.

Eros has also been smoothed and rounded by glancing blows during its catastrophic past. This cosmic sculpture rivals the smaller bronze and marble forms of Constantin Brancusi (1876–1957) and Henry Moore (1898–1986). Equally beautiful is the broad, curved depression that connects two mountains on Eros, each thousands of meters high (Figs. 12.14, 12.15).

Far from being a barren lump of rock, Eros has a dusty, boulder-strewn landscape. Despite its weak gravity, the diminutive asteroid has managed to hold on to about 7000 boulders larger than 15 meters, forced out of craters and pulled back to the surface during the relentless bombardment of its past. Some of the isolated stones are as large as a house, and up to 100 meters across.

The positions of the rocks on Eros indicate that at least 3000 of them were ejected from a single impact crater, perhaps a billion years ago. Some of these boulders went straight up and straight down. Most of the remainder traveled as far as two-thirds of the way around the asteroid, in all directions, before finally coming to rest on the surface.

Smaller rocks and a loose layer of dirty debris came into view when the *NEAR Shoemaker* spacecraft moved in to land on the boulder-strewn surface of 433 Eros. It took pictures as close as 250 meters above the surface, showing features as small as a golf ball (Fig. 12.16). They indicate that Eros is something between a very big rock and a planet, large enough to hold on to its pieces yet too small to lose its odd, distorted shape.

Asteroid 433 Eros has a bulk mass density of 2670 kilograms per cubic meter, comparable to that of terrestrial rocks, and Eros is just like an extra big, chipped rock. Millions of range, or distance, measurements with a laser aboard *NEAR Shoemaker* have established the asteroid's complex shape and topography. Comparisons of this

Fig. 12.14 Back in the saddle again This image of the saddle region on 433 Eros was taken on 22 March 2000 by the *NEAR Shoemaker* spacecraft. The saddle is about 10 kilometers across. It may be the scar of an ancient crater, or somehow related to a different large crater on the opposite side of the asteroid. (Courtesy of NASA/JHUAPL.)

Fig. 12.15 The south saddle of Eros This mosaic of four images, photographed on 26 September 2000, was taken as the *NEAR Shoemaker* spacecraft looked down on the saddle region of the asteroid 433 Eros. A broad, curved depression stretches vertically across the image, as if scooped out by a cosmic sculptor. A boulder-rich area is seen in the lower right. (Courtesy of NASA/JHUAPL.)

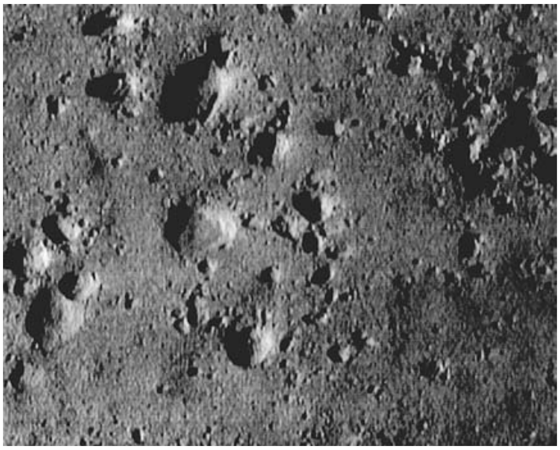


Fig. 12.16 Close-up view of the surface of 433 Eros This *NEAR Shoemaker* picture of the surface of asteroid 433 Eros was taken from a range of 250 meters on 12 February 2001, just before landing on the asteroid. The image is just 12 meters across, and the cluster of rocks at the upper right measures 1.4 meters across. (Courtesy of NASA/JHUAPL.)

shape with radio-tracking of the gravitational pull on the orbiting spacecraft show that the mass density of Eros' interior must be nearly uniform. Thus, the asteroid 433 Eros is mostly solid throughout, with a uniform composition and homogeneous internal structure, rather than a loosely bound collection of smaller components. It is an undifferentiated object that never separated into a distinct crust, mantle and core. Although 433 Eros has been heavily fractured and whittled away by impact collisions, it was never completely demolished and reformed as a rubble pile.

Most asteroids are probably covered with a blanket of dust, pebbles and rocks that rests on solid bedrock. This layer of loose rock particles is known as the *regolith*, from the Greek word for "rock layer". The regolith formed during repeated bombardment by small meteorites that break apart the surface rock. The Earth's Moon and Mars contain such dusty surface layers. Small asteroids may have a thin coating of dust, while the largest asteroids could have a thicker, powdery veneer, even thousands of meters deep.

The fine, dusty material on Eros has settled downhill, collecting in ponds that are tens of meters wide and a few meters deep. Hundreds of them are found in low-lying hollows, such as the bottoms of craters. The powdered deposits have flat, level surfaces, resembling ponds of water on Earth, but there is no water on a rocky asteroid like Eros, and there hasn't been any for billions of years. Something else is causing the mobile soil to flow down hill. Since the ponds are found in well-lit areas, scientists speculate that the dust moves when it has been in the Sun too long, but the details of the sorting mechanism are not well understood.

Hayabusa orbits 25143 Itokawa, a rubble pile

On 9 May 2003 the Japanese Institute for Space and Astronautical Sciences (ISAS) launched its *MUSES-C* spacecraft to the near-Earth asteroid 25143 Itokawa. As is the custom, the spacecraft was given another name following successful launch; it is now known as *Hayabusa* – Japanese for "falcon".

Propelled by an ion-drive engine, *Hayabusa* arrived at asteroid Itokawa on 30 September 2005, hovering at an altitude of just 7 kilometers above its surface for about a month. The falcon then landed on the asteroid surface for brief 30-minute stays on 19 and 25 November, like a bird swooping down for fish near the surface of the sea.

Asteroid 25143 Itokawa is a small rocky body, just 535 meters in longest dimension, spinning a couple of times a day and belonging to the S-type of spectroscopic class common in the inner portion of the asteroid belt. It is a nearby Apollo asteroid, discovered in 1998, which follows a 1.5-year orbital path a distance of 0.953 AU from the Sun, just inside the Earth's orbit, to 1.695 AU, just outside the orbit of Mars. The asteroid was thus selected as one of the most accessible targets for the low-energy launch vehicle, as well as providing an opportunity to learn more about small asteroids that regularly move past the Earth and may one day collide with it.

The *Hayabusa* images indicate that 25143 Itokawa has a lumpy, oblong shape that has been likened to a sea otter. It is composed of two parts: a smaller "head" and a larger "body" with a smooth "neck" separating them (Fig. 12.17). The head and body could have been formed separately and later came into contact at a slow relative speed, or the depressed neck region might have been formed by a large impact with a single entity. Shaking and jostling of the asteroid by former impacts may account for the absence of craters, filling them in, as well as the sorting of large boulders and small gravel into different places on the surface.

This asteroid seems to be nothing but pieces, hanging together by its weak gravity, continuously pulverized by impacts and sorted by vibrational shaking. Impact craters seem to be filled as quickly as they form, and smooth, deep gravel beds have been created. Itokawa's vital statistics reflect its small size, with a mass of just $(3.58 \pm 0.18) \times 10^{10}$ kilograms and a low bulk mass density of

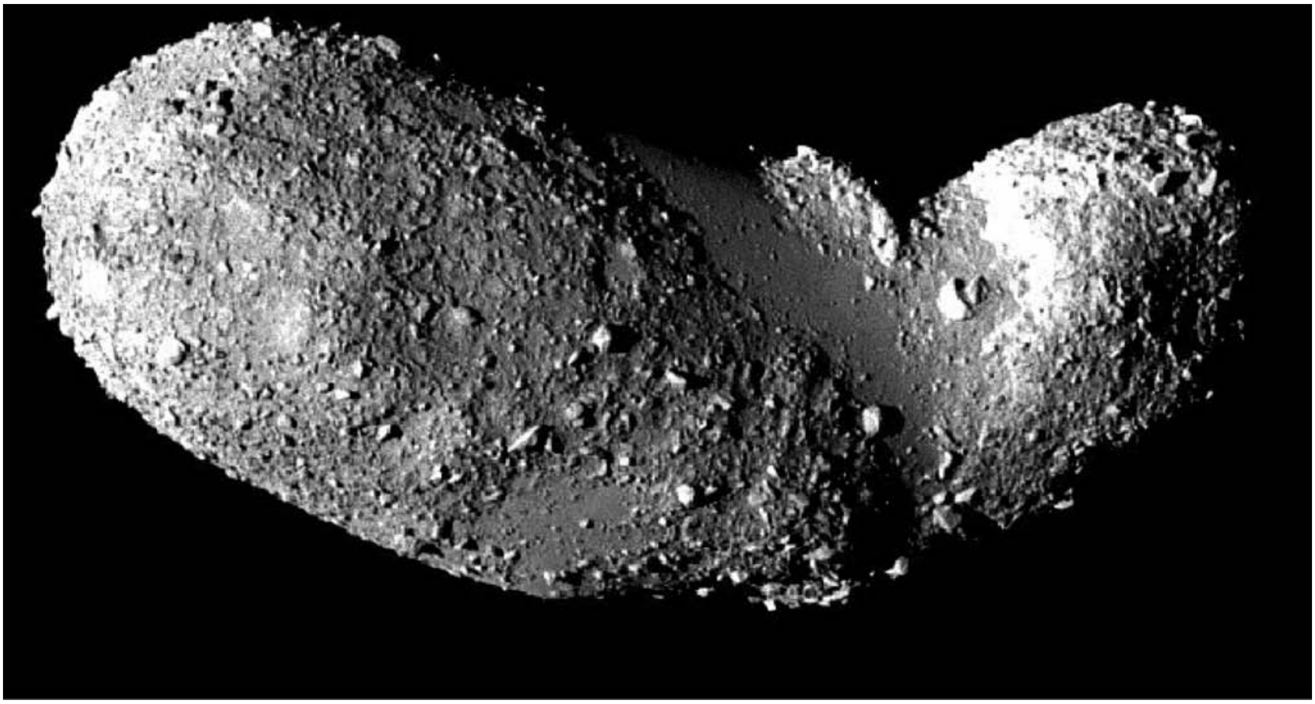


Fig. 12.17 Asteroid 25143 Itokawa, a rubble pile The small asteroid 25143 Itokawa is just 500 meters long, and it has an orbit that crosses that of the Earth. This image, taken from the Japanese spacecraft *Hayabusa* in late 2005, strongly suggests that the asteroid is a loose pile of rubble rather than a solid rock. The difference between the smooth and rough terrain might be due to impacts that jostled and shook the asteroid, creating a separation of large and small rocks on the surface. (Courtesy of ISAS/JAXA.)

1950 ± 140 kilograms per cubic meter, indicating a porous interior filled with holes something like sand.

Thus 25143 Itokawa is considered to be a veritable rubble pile because of its low bulk density, high porosity, boulder-rich appearance, shape and absence of craters. It may have been formed by the early collision and breakup of a pre-existing parent asteroid followed by re-agglomeration into a rubble-pile.

Hayabusa returned to Earth from its seven-year, six-billion-kilometer trip on 13 June 2010, burning up in the atmosphere after releasing its small asteroid return capsule that was parachuted to the vast Australian Outback, where it was recovered and transported to Japan.

On 15 April 2010, President Barack Obama (1961–) announced a goal of sending astronauts beyond the Moon and into deep space, to encounter and perhaps land on an asteroid by 2025. The astronaut mission may develop the capacity to mine fuel and metals from an asteroid, as a precursor to a human mission to Mars, or to deflect a near-Earth asteroid headed for collision with Earth. It reminds one of Antoine de Saint-Exupéry's (1900–1944) charming novella about *The Little Prince*, who lives on an asteroid the size of a house and named B-612, making profound observations about life and human nature.

The incredible successes of the *NEAR Shoemaker* and *Hayabusa* missions have provided background data for the *Dawn* mission that will rendezvous with two of the largest bodies in the asteroid belt, 4 Vesta and 1 Ceres, beginning in 2011 and 2015, respectively (Focus 12.2).

12.5 Meteorites

Space rocks

Most meteors, or shooting stars, are produced by tiny fragments of comets: icy dust and pebbles that burn up in the air and never reach the ground, but occasionally a stone will fall from the sky, producing a brilliant trail of light flashing across the night sky. A rumbling sound and what appears to be a great burst of sparks may accompany it. These are fireballs and they are produced by tougher chunks of matter from space, resembling rocks (Fig. 12.18).

Extraterrestrial chunks of rock and metal that survive the fiery descent through the atmosphere and reach the ground have been given the name meteorites. Strictly speaking a *meteoroid* is the solid object in space that appears as a *meteor* when it lights up in the atmosphere and becomes a *meteorite* if it reaches the ground.

Focus 12.2 The *Dawn* mission to asteroids Ceres and Vesta

The *Dawn* spacecraft began a 3-billion kilometer voyage to the main asteroid belt on 27 September 2007, acquiring a gravity assist during an encounter with Mars in February 2009. The mission will be the first to orbit two solar-system bodies in a single voyage, revolving about the asteroid 4 Vesta from September 2011 to April 2012 and orbiting the dwarf planet–asteroid Ceres in 2015. By utilizing the same set of instruments, it can make accurate comparisons of two very different objects, including their mass, bulk density, shape, surface topography, tectonic history, and elemental and mineral composition.

Vesta is a dry, rocky, evolved and differentiated body with basaltic lava flows on its surface, perhaps resembling the magma maria on the Earth's Moon. The asteroid is representative of the building blocks that constructed the terrestrial planets of the inner solar system. The other object, 1 Ceres, has a surface with water-bearing minerals and may well be icy on the outside with water on the inside. It may be representative of the building blocks of the outer planets.

Dawn is expected to answer questions about the formation of the solar system by studying two contrasting protoplanets, one dry and the other wet, that never gathered together with other objects to form a planet. According to current theories, their differing properties are the result of their being formed and evolving in different parts of the solar system, and they are expected to characterize conditions and processes at the earliest moments – the dawn – of the planets.

Dawn is expected to arrive at 1 Ceres just five months prior to the arrival of the *New Horizons* spacecraft at Pluto in the cold outer precincts of the planetary system.

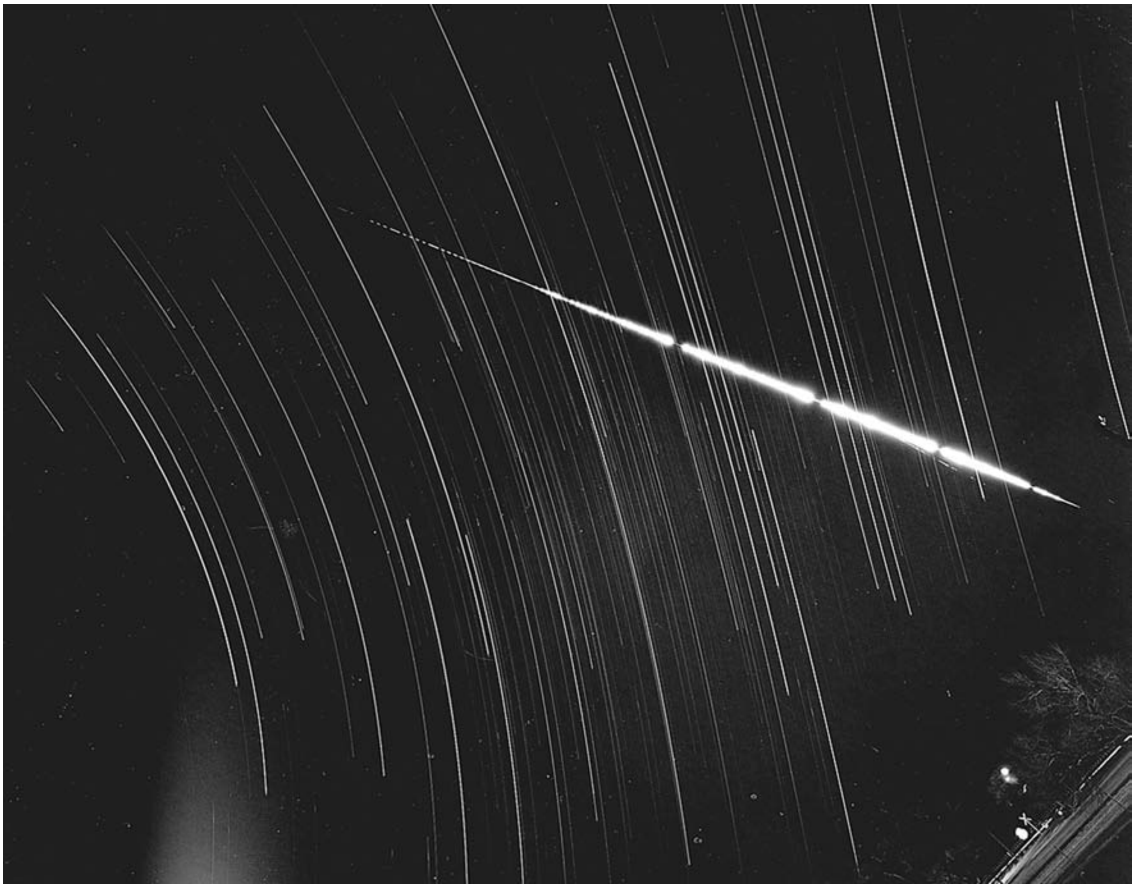


Fig. 12.18 Fireball A great flash of light, called a fireball, is produced when a large meteoroid streaks through the atmosphere. It is often accompanied by sonic booms and rumbling noises. A camera-chopping shutter used for timing and velocity determinations produced spaces between the luminous segments of the fireball's trajectory. The faint curved lines in the background are star trails caused by the Earth's rotation during the three-hour exposure. (Courtesy of SAU.)

Fig. 12.19 Antarctica The midnight Sun illuminates the wind-swept ice at the bottom of the world. Numerous meteorites have been found embedded in the ice in this region near Allan Hills, Antarctica. These meteorites are probably fragments of asteroids that once had orbits between those of Mars and Jupiter, but a few of them may have come from the Moon or even Mars. (Courtesy of Ursula Marvin/CfA.)

Meteorites have long been recognized as celestial objects. The Acts of Apostles in the New Testament (Acts 19:35) refers to a temple dedicated to Artemis in which there is a "sacred stone that fell from the sky". These black objects have also been found in the Egyptian pyramids with a hieroglyph meaning "heavenly iron".

Until relatively recently, only about 10 meteorites were recovered each year. They were ones that happened to fall near populated regions of the Earth, and many more must be buried at the bottom of the ocean, lost in the jungles, or buried in desert sand.

The Antarctica lode

In 1969, a group of Japanese scientists discovered a bountiful source of meteorites on the blue ice fields at the bottom of the world (Fig. 12.19), leading to a dramatic increase in the number of recovered meteorites. During the next four decades, tens of thousands of these cosmic rocks have been harvested from the Antarctic ice. The most productive areas were near the Allan Hills in Victoria Land and the Yamato Mountains in Queen Maud Land.

Many of the objects recovered from Antarctica must be fragments of the same meteorite, but they also represent thousands of other different meteorites, more than doubling the number found on Earth. Before scientists traveled to Antarctica, the world's meteorite collections only had about 2600 different specimens, most of them collected during the past two centuries.

A meteorite landing on the Antarctic ice becomes buried in compressed snow, and is quickly frozen into the thickening ice. The cosmic rock soon sinks to great depths, where it remains preserved against corrosion. The tremendous mass of the ice, which reaches a thickness of 4 kilometers, squeezes the ice downward and pushes it outward toward the edges of the continent. The buried meteorite becomes caught in these huge ice-flows that move like rivers under the surface, creeping along at rates of several meters per year.

Fig. 12.20 Stony meteorite Fragment of the stony, chondrite meteorite that fell near Johnstown, Colorado. (Courtesy of the American Museum of Natural History.)

Some of the ice and its enclosed meteorites ultimately reaches the sea and breaks off as icebergs. At other locations the ice-flow encounters an obstacle, such as a mountain range. The moving ice-flow then thrusts the meteorite upward and forces the buried rock to the surface. Strong winds corrode and wear away the surface, removing the ice covering and exposing the meteorite. The dark crusts of the meteorites are then easy to spot against the bright icy background.

Meteorites recovered in Antarctica have stayed there for relatively long times. Radioactive dating indicates that most of them have spent about half a million years entombed in the ice, remaining virtually unchanged since the time they struck the Earth. By way of comparison, most of the meteorites found elsewhere on Earth have fallen within the past 200 years.

Typical meteorites

Meteorites, together with rocks returned by astronauts from the Moon and grains of dust collected from the high levels of the Earth's atmosphere or from comet Wild 2 by the *Stardust* mission, are the only samples we know of extraterrestrial material. Although a meteorite's surface is usually coated with dark glassy material, known as a *fusion crust*, that melted during its descent through the atmosphere, the heat of friction did not have time to penetrate deeply into the falling rock. So most recovered meteorites are cold inside, and their interiors are unaffected by their rapid fall though the atmosphere. Meteorites may therefore be cut open and examined with microscopes and subjected to chemical analysis that reveals their original constitution.

Most meteorites that have been seen to fall and then recovered are stones rather than chunks of metal (Fig. 12.20). About 94 percent of the fallen meteorites are stones, 5 percent irons and 1 percent stony-irons (Table 12.3). About 90 percent of the stony meteorites are, in turn, classified as chondrites. So most of the meteorites that fall to Earth are chondrites, which contain millimeter-sized chondrules (Fig. 12.21). The name "chondrite" is derived from the ancient Greek word *chondros*, meaning "grain" or "seed". The other 10 percent of the stony meteorites are achondrites, which show signs of past igneous activity (Fig. 12.22).

The chondrites have been additionally divided into clans with common properties, suggesting that each clan formed in the same region of the solar system. They are the ordinary chondrites, which are the most abundant, the carbonaceous chondrites, and the enstatite

Table 12.3 Classes of fallen meteorites

Name	Composition	Density[a] ($kg\,m^{-3}$)	Percent[b]
Stones	Silicates (75 to 100 percent) and metallic nickel-iron (0 to 25 percent)	3500 to 3800	94
Irons	Nickel, iron (100 percent)	7600 to 7900	5
Stony-irons	Silicates (50 percent) and metallic nickel-iron (50 percent)	4700	1

[a] For comparison, typical rocks on Earth are largely silicates with densities in the range 3100 to 3300 kilograms per cubic meter ($kg\,m^{-3}$), so meteorites are usually denser than other rocks found on the Earth's surface.
[b] Percent of total meteorites.

chondrites, named for their high abundance of the mineral enstatite.

Most meteorites are denser than terrestrial rock. So, if you find a dark, rather smooth rock that you suspect of being a meteorite, it must weigh at least as much as an ordinary Earth-born rock of the same volume if it is to pass muster as a rock from the sky.

But, of course, every rule has an exception – except that rule. There is a rare class of meteorites, with the ponderous name carbonaceous chondrites, that are fragile and have unusually low densities in the range 2200 to 2900 kilograms per cubic meter, making them less dense than an average terrestrial rock. They often contain appreciable amounts of carbon and water and they are considered to be among the most primitive and least altered samples of solids in the solar system.

Chronology of the meteorites

If we ask "How old is a meteorite?" the question can have several meanings. Each meaning refers to the time since a significant event in the history of a meteorite, and there are three of them: the formation, breakup and impact of a meteorite with Earth.

1 Formation

Meteorites are as old as the solar system, dating back to its earliest days. Most of them were formed at about the same time as the planets. They accumulated directly from the primeval solar nebula and they have compositions similar to that of the Sun, except for their lack of hydrogen and helium.

The dates of formation of the meteorites can be determined by radioactive dating, in much the same way that the ages of lunar rocks were determined. The relative concentrations of the decay products of elements such as rubidium and uranium reveal the time since these rocks were formed. Such measurements indicate that the carbonaceous chondrite meteorites formed 4.566 ± 0.002 billion years ago, and that some stony and iron meteorites are 4.55 ± 0.07 billion years old.

Rounding off the numbers, and allowing for possible systematic errors, we obtain an age of about 4.6 billion years for most meteorites. That is when the mineral grains in the meteorites crystallized, and their radioactive clocks started ticking. The presence of decay products from short-lived radioactive elements additionally indicates that the meteorites formed in just a few million years some 4.6 billion years ago.

The meteorites are the oldest rocks that we can touch. They are hundreds of millions of years older than the oldest rocks on the surface of the Earth. All of the planets and satellites are thought to have originated together with the meteorites 4.6 billion years ago, but erosion and geological processes have destroyed the original rocks on Earth. As a result, it is meteorites that reveal the age of the solar system and give clues to its origin.

2 Breakup and exposure

Another type of radioactivity also occurs in meteorites — radioactivity that is continually being caused by cosmic rays in the solar system. These "rays" are not rays in the usual sense; they are very energetic protons that bombard the meteorites and penetrate their surface for short distances. This cosmic-ray bombardment performs a bit of alchemy and transforms some of the atoms of the meteorite into radioactive nuclei. These radioactive nuclei slowly disintegrate, creating "daughter" nuclei. As the meteorite continues to be exposed to cosmic rays, the daughter nuclei

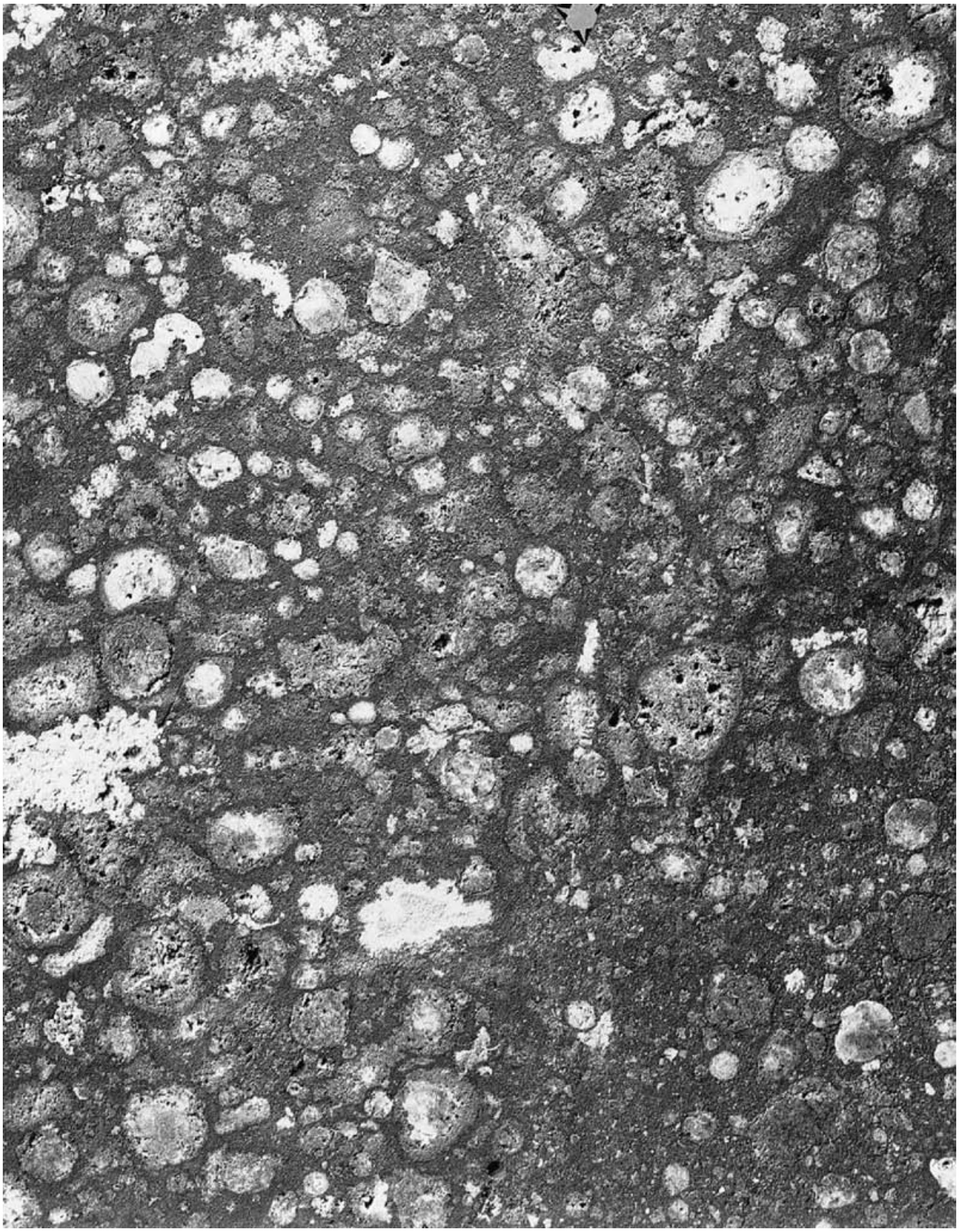


Fig. 12.21 Chondrules in Allende This photomicrograph of a thin section of the Allende meteorite shows numerous round silicate chondrules together with irregular inclusions. The meteorite section is 0.021 meters across and 0.027 meters high. (Courtesy of the Smithsonian Institution.)

become more and more abundant, and by a careful measurement of the amount of daughter nuclei in a meteorite it is possible to estimate the duration of this exposure interval, or the exposure age of the meteorite.

Now, the exposure ages of the meteorites that have been recovered on Earth are remarkably short in astronomical terms. Typically they are between 5 and 60 million years – just an instant in the life of the solar system and the meteorites. Evidently the meteorites have spent most of their lives shielded from cosmic rays. Astronomers now believe that most meteorites spent a larger portion of their life protected inside a parent body that was much smaller than the Earth but larger than a typical meteorite. According to this view, an important event in the chronology of most meteorites was the breakup of a parent body, exposing smaller fragments to space and to the bombardment by cosmic rays. The exposure ages measure the time that has elapsed since the breakup took place.

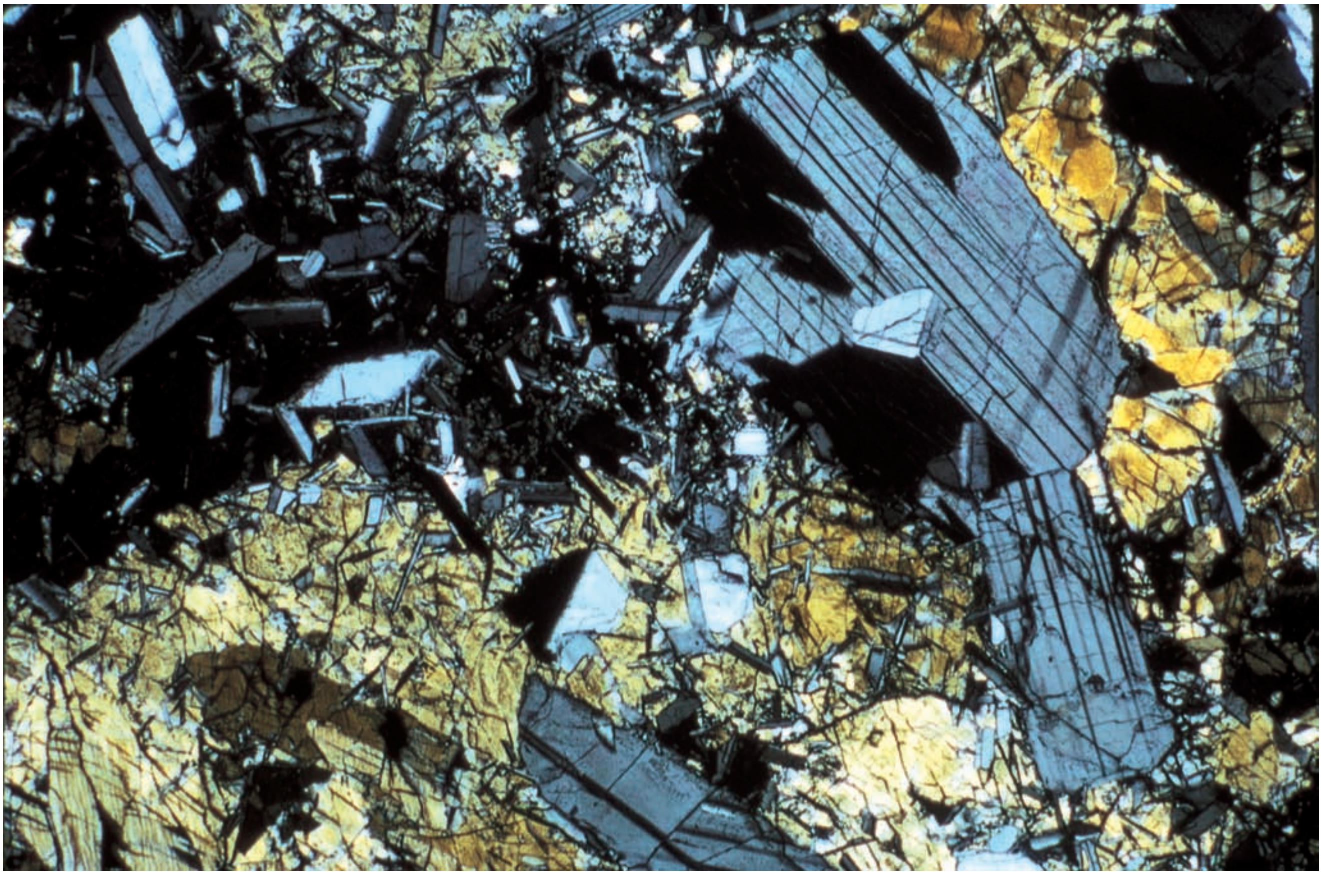


Fig. 12.22 Achondrite meteorite A photomicrograph of the achondrite meteorite that fell near Juvinas, France, on 15 June 1821. It contains basalt, resulting from the melting and separation of material inside an asteroid-size parent body. The section shown here is 0.0032 meters across. (Courtesy of Martin Prinz, American Museum of Natural History.)

3 Collision with the Earth

Finally, when a meteorite falls through the blanket of the Earth's atmosphere, it becomes protected from cosmic-ray bombardment. No more radioactive atoms are created, and the ones that already exist inside the meteorite begin to decay, like the slow ticking of a clock.

In this way, the atoms of a meteorite carry a record of their chronology that can be unlocked with radiochemistry.

Rare and exotic finds

The frozen cargo of the Antarctica ice includes at least a dozen, greenish-brown meteorites that are strikingly similar to the welded highland rocks from the Earth's Moon. The abundance of various elements and gases in these meteorites are virtually identical to those found in lunar rocks; at the same time, they are unlike those found in any other known meteorite or terrestrial rock. These small stones were apparently blasted off the Moon by impacting objects.

Out of the thousands of stony meteorites now found in terrestrial collections, roughly a dozen are believed to be pieces of Mars. They were similarly blasted into space by impacting objects, with such force that they escaped the red planet's gravitational pull and eventually reached Earth. One of them, dubbed ALH (for Allan Hills in Antarctica) 84001 was once thought to contain evidence for ancient microscopic life on Mars, but that evidence has subsequently been attributed to other causes.

The meteorites from Mars contain small amounts of water and water-altered minerals. Moreover, gases trapped in bubbles within some of them have the exact, unique composition as the gases in the Martian atmosphere, as measured by the *Viking* landers, and found nowhere else. So there can be little doubt that these meteorites came from Mars.

They are often referred to as the SNC, pronounced "snick", meteorites, short for Shergotty (India), Nakhla (Egypt), and Chassigny (France) – three locations where they were observed to fall from the sky. Radioactive dating indicates that many of the SNCs were molten in the relatively recent history of our solar system. Those named for Nakhla and Chassigny hardened into solid rock 1.3 billion years ago, while the ones found near Shergotty solidified from molten lava just 180 million years ago. In comparison, most other meteorites solidified from molten materials between 4.5 and 4.6 billion years ago.

Organic matter in meteorites

For more than a century, organic molecules have been suspected inside meteorites, although their presence in newly arrived meteorites led to the suggestion that they were the result of terrestrial contamination. But their existence became a certainty when 20 kinds of amino acids were found in the carbonaceous chondrite meteorites from Antarctica. These meteorites had lain in a sterile environment and were collected using sterile procedures. The organic matter in these carbonaceous chondrites apparently formed within the meteorite parent bodies, probably when water was bound into the structure of its clay minerals.

Many organic compounds, such as the amino acids, can come in two versions that are mirror images of each other. They are identified as left- and right-handed, based on their ability to rotate light in one direction or another. All living organisms on Earth use only left-handed amino acids. In contrast, the carbonaceous chondrites contain roughly equal amounts of both types, including the right-handed amino acids that are not found in living systems on Earth. This provides convincing evidence that the organic matter in meteorites did not originate on our planet, and that it is not directly related to life as we know it.

This discovery implies that the amino acids and other organic molecules probably existed in the solar system a billion years before the appearance of living things on Earth. But the molecules found in carbonaceous chondrites are generally thought to be of non-biological origin. So the organic molecules found in meteorites are not in themselves vestiges of extraterrestrial life. But they are certainly primitive, and their cousins may have been the precursors to living matter.

The asteroid–meteorite connection

What is the source of meteorites? The primitive nature of most meteorites, called chondrites, indicates that they came from objects that have not experienced geologic processes. Most of them are not igneous, and did not go through a hot, liquid stage. This is most easily understood if their parent bodies were small asteroids. Only a small fraction of meteorites, known as the achondrites, formed by igneous processes in larger parent bodies, most likely one of the original asteroids.

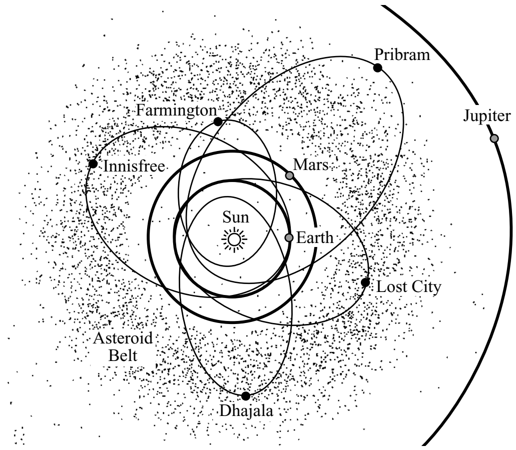


Fig. 12.23 Meteorite orbits The calculated orbits of five meteorites, inferred from their trajectory before hitting the ground. All of them originated in the asteroid belt, indicating that these meteorites are chips off asteroids.

There is little doubt that most of the meteorites have come from the asteroid belt. They are probably chips off wayward asteroids, and there are three pieces of evidence for this conclusion. They are the orbits, colors and crystalline structure of meteorites.

1 Orbits

Photography of meteorites as they descend through the Earth's atmosphere can be used to determine their precise speed and direction of motion when they encountered the Earth. From these data, their orbits may be inferred, and many of the objects came from space beyond Mars, in the main belt of asteroids (Fig. 12.23).

2 Colors

The surface composition of asteroids has been inferred by breaking down their reflected sunlight into its component colors. Such spectral displays are similar to those of meteorites, suggesting that the meteorites are the debris of colliding asteroids. The relative abundance of asteroid types are not like those of the fallen meteorites, but this may simply reflect the ease or difficulty in sending asteroid fragments to Earth.

The sunlight reflected from C-type asteroids, that far outnumber all other asteroid types, closely resembles that of the relatively scarce carbonaceous chondrite meteorites. But the C-type asteroids reside in the remote, outer part of the main belt, furthest away from the Earth. The most common meteorites, the ordinary chondrites, display spectral colors that resemble those of the S-type asteroids found in the inner half of the asteroid belt. Due to its proximity, most meteorites are expected to originate from this part of the main belt. The light reflected from rare,

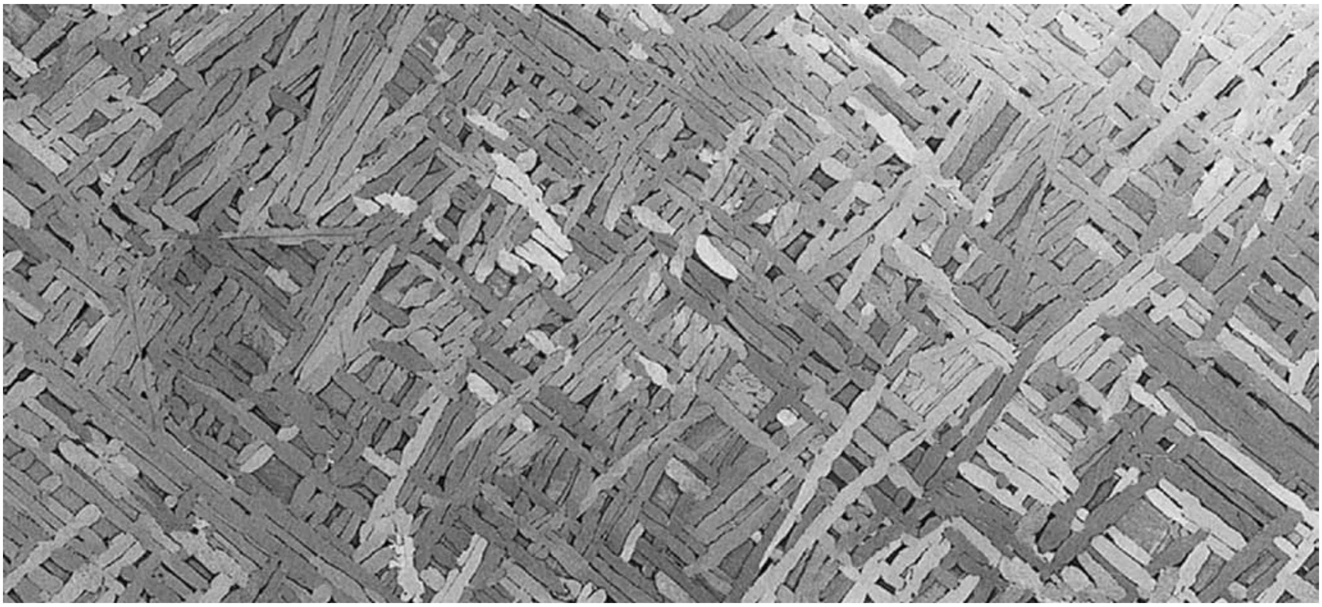


Fig. 12.24 Widmanstätten pattern When polished and etched with acid, an iron meteorite displays this distinctive Widmanstätten pattern produced by crystals of two different iron–nickel alloys. The pattern provides evidence that this meteorite was once buried within a parent body between 50 and 200 kilometers in radius. This sliced specimen is about 0.05 meters across. (Courtesy of the Smithsonian Institution.)

bright, M asteroids, which reside near the middle of the main belt, matches the spectrum of relatively scarce iron meteorites.

Nevertheless, the color coordination is not exact. The S-type asteroids have redder overall colors and subdued light absorption compared with the abundant meteorites. This discrepancy may be explained if solar radiation or small impacting cosmic particles gradually altered the thin outer layers of asteroids, darkening and reddening their surfaces. Similar "space weathering" apparently makes the lunar surface much redder than the color of unexposed rocks returned from the Moon. And the *Galileo* spacecraft showed that material near the sharp-edged, relatively young craters on 951 Gaspra and 243 Ida is slightly bluish, while the low-lying, older areas are slightly reddish. So the evidence now suggests that the color of freshly exposed asteroid surfaces gradually redden with time. The bluish surfaces are recently exposed, while the red areas have been weathered for millions and even billions of years.

Powerful collisions most likely excavate meteorites from deeper layers inside asteroids, which retain their pristine, unweathered color. So the surface appearances of asteroids are slightly misleading. Deep down inside, the most common, inner main-belt asteroids are probably very similar to the most common meteorites.

Instruments aboard *NEAR Shoemaker* revealed that 433 Eros, also an S-type asteroid, has the same basic composition as some ordinary-chondrite meteorites. The composition of the surface of asteroid 1 Ceres is similar to that of a water-rich carbonaceous chondrite meteorite. Asteroid 4 Vesta is an exception, bearing signs of significant heating, differentiation and surface lava flow. Five percent of the found meteoritic samples on Earth, the Howardite Eucrite Diogenite (HED), meteorites, are thought to be the result of a collision or collisions with Vesta.

3 Crystalline structure

When the majority of iron meteorites are cut and polished and then are etched with acid, a delicate and complex pattern emerges (Fig. 12.24). It is produced by regions of crystalline structure, depending on the local orientation of the crystals in the iron. The sizes and shapes of these crystals indicate that they grew very slowly, and that the meteorite must have been hot, almost to the melting point, for tens of million of years. It probably cooled at the rate of a few degrees in a million years.

Such a slow cooling rate is compelling evidence that the meteorites were once inside a sizeable parent body. If a small meteorite had been exposed to space when it was still hot, it would have cooled in a matter of days. Small meteorites cool rapidly because their material is close to the surface, through which the heat can escape. Large crystal patterns would not have grown in such small bodies.

The meteorites that retained their heat for millions of years must have been buried within parent bodies between 50 and 300 kilometers in radius, and this is just the size of large asteroids.

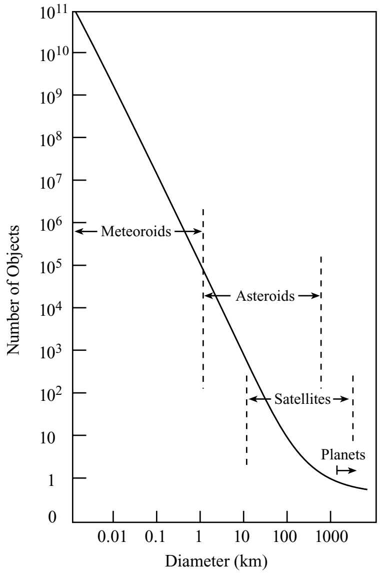


Fig. 12.25 Interplanetary objects Repetitive collisions between interplanetary objects have produced many more small meteoroids than large ones. Some of the largest asteroids are comparable in size to small moons, and ongoing collisions between asteroids have produced numerous smaller meteorites.

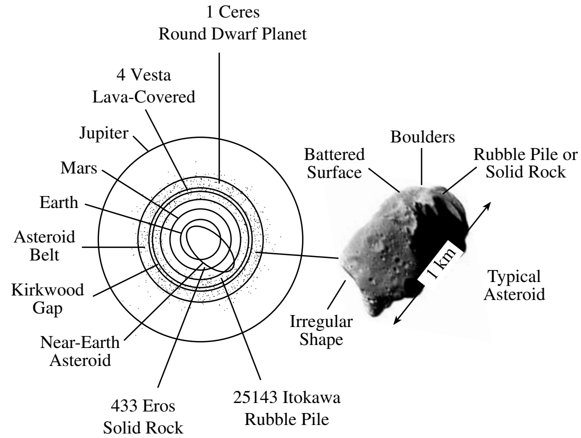


Fig. 12.26 Summary diagram

Additional suggestive evidence is the relationship among the sizes of asteroids, meteorites, and meteoroids (Fig. 12.25). The classes are not mutually exclusive, and there is considerable overlap. The simplest explanation of all the evidence is that meteorites are the debris of collisions among the asteroids.

Whatever their precise history, the asteroids and meteorites are primitive objects that can act as beacons to the past. They represent a tableau of the ancient objects that littered space 4.6 billion years ago, and they carry messages for our understanding of the formation of the solar system.

13 Colliding worlds

- At least 20 pieces of a comet hit Jupiter on 7 July 1992, producing explosive fireworks and dark scars that fascinated astronomers throughout the world.
- Some comets are on suicide missions to the Sun, diving into our star and being consumed by it.
- Most of the impact craters on the Earth disappeared long ago, but a few of the relatively recent ones have been located from space.
- An asteroid wiped out the dinosaurs when it hit the Earth 65 million years ago.
- If an asteroid or comet of about 10 kilometers in size hit the Earth, the horrific blast could generate overpowering ocean waves, block out the Sun's light and heat, ignite global wildfires, drench the land and sea with acid rain, and produce deadly volcanoes on the other side of the Earth.
- The Earth is immersed within a cosmic shooting gallery of potentially lethal, Earth-approaching asteroids that could collide with our planet and end civilization as we know it.
- The lifetime risk of your dying as the result of an asteroid striking the Earth is about the same as death from an airplane crash, but a lot more people would die with you during the cosmic impact.
- It is estimated that the Earth receives a direct hit by an asteroid about two kilometers in size every million years or so, resulting in a global catastrophe. It could happen tomorrow or it might not occur for hundreds of thousands of years.
- Astronomers are now taking a census of most of the Near-Earth Objects that are big enough and close enough to threaten us with global destruction.
- With enough warning time, we could redirect the course of an asteroid or comet that is headed for collision with the Earth.

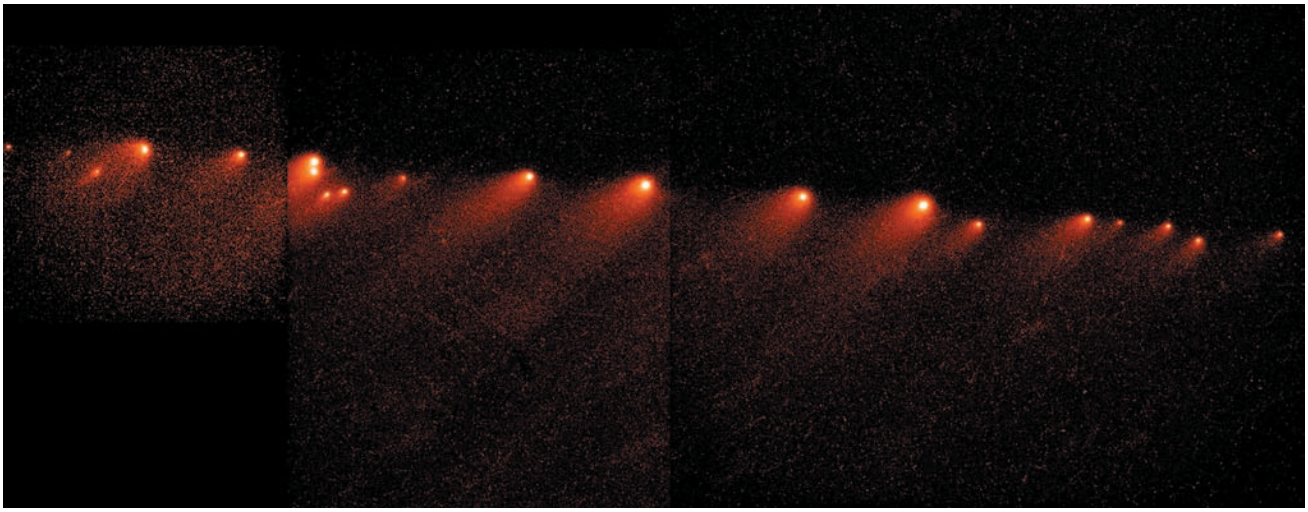


Fig. 13.1 Fragments of comet Shoemaker-Levy 9 In July 1992, this comet passed so close to Jupiter that the icy material of its nucleus was torn apart by the differential gravitational forces of the giant planet. This panoramic image of the comet fragments was taken from the *Hubble Space Telescope* in January 1994, eight months after they were discovered and six months before they dived into the atmosphere of Jupiter. The length of the string of comet pieces is about 1.1 million kilometers, three times the distance from the Earth to the Moon. The largest of the fragments in the string is about two kilometers across. Each fragment mimics a larger comet with a round coma and a dusty tail. (Courtesy of NASA/STScI.)

13.1 A comet hits Jupiter

Comet Shoemaker-Levy 9

Eugene Shoemaker (1928–1997), his wife Carolyn, and the amateur astronomer, David Levy (1948–) were involved in routine observations the cloudy evening of 23 March 1993. They were continuing a ten-year search for comets and asteroids that might be headed toward the Earth using the small 0.46-meter (18-inch) wide-field photographic telescope at Palomar Observatory in California. Two days later, when Carolyn examined the images, taken on fogged film, she saw an elongated feature that looked to her like a "squashed comet". When the discovery was confirmed with better telescopes, the stretched-out blur of comet light was resolved into several little comets aligned along a single straight line projected in the sky, each with a nearly spherical coma and elongated dust tail (Fig. 13.1). In accordance with tradition, it was named comet Shoemaker-Levy 9, after the last names of the discoverers, ninth in a series of objects the trio found traveling around the Sun in short-period orbits.

Comet Shoemaker-Levy 9 (SL9) consisted of the pieces of a former single comet that had been trapped in a two-year orbit around Jupiter for decades. But when it traveled too close to Jupiter in 1992, the comet was torn apart by the difference between the planet's gravitational attraction on the near and far side of the comet, ripping the fragile comet into at least 20 observable pieces.

A few previous comets had been known to orbit Jupiter temporarily, and the disruption wasn't unprecedented. What made SL9 unique was that the broken comet was inexorably hurtling toward a collision with Jupiter. Orbital calculations indicated that the train of comet fragments would plunge into the giant planet in July 1994, two years after the former single comet's disruption and more than one year after the discovery of its pieces (Fig. 13.2).

Such advance knowledge of a collision with any planet was unprecedented in human history. Numerous craters on the terrestrial planets, as well as the Earth's Moon and some of Jupiter's satellites, bear witness to cosmic collisions of the past, but now for the first time astronomers could see the collision happening before their very eyes. The anticipated impact was an incredible opportunity, occurring just once in the lifetime of any astronomer and most likely once in a millennium.

Impact of a comet with Jupiter

The collision of comet Shoemaker-Levy 9 with Jupiter in July 1994 was perhaps the most widely witnessed event in astronomical history. Practically every telescope in the world was trained on Jupiter during impact week, between 16 and 22 July 1994. Infrared heat detectors were placed at the focal point of the Keck Observatory's giant 10-meter telescope atop Mauna Kea in Hawaii, and the *Hubble Space Telescope* was poised to record the event at visual wavelengths. Every other major astronomical observatory

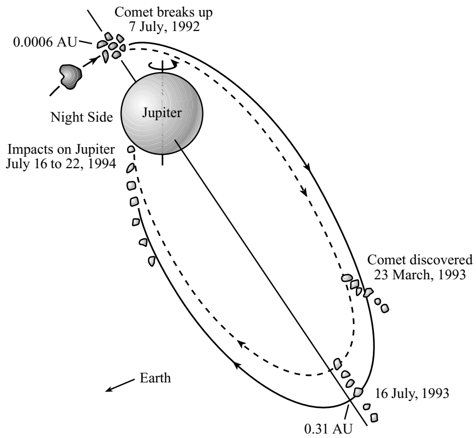


Fig. 13.2 Final orbit of comet Shoemaker-Levy 9 This comet was orbiting Jupiter for more than half a century, until it was ripped apart during a close encounter with the planet and collided with it two years later. The disruption occurred on 7 July 1992 when the comet passed within 0.0006 AU, or 90 000 kilometers, from the planet's center. Since Jupiter has a radius of just over 70 000 kilometers, the comet passed within about 20 000 kilometers of the planet's cloud tops. Jupiter's unequal gravitational pull on the near and far sides of the comet nucleus then tore the object apart. Carolyn and Eugene Shoemaker and David Levy discovered the comet fragments on 23 March 1993, when the broken comet was almost at its farthest distance from Jupiter, at 0.31 AU. One by one the icy fragments exploded in Jupiter's cloud tops during impact week, from 16 to 22 July 1994.

Fig. 13.3 Dark impact scars of comet fragments hitting Jupiter This *Hubble Space Telescope* image shows several dark spots (*right*) on Jupiter, each marking the impact site of a fragment of Comet Shoemaker-Levy 9. The Earth-sized scars remained visible for about five months, until the winds in Jupiter's outer atmosphere pulled them apart. A thin expanding ring of dark material, suggesting waves spreading out from the impact explosion, surrounded some of the dark central spots. Jupiter's Great Red Spot is also prominent, as it has been for centuries. (Courtesy of NASA/STScI.)

participated, as did numerous amateur astronomers from their own backyards.

Detailed calculations indicated that the collisions would be on the dark "back" side of Jupiter, hidden from the Earth's view by the body of the giant planet. So it might be something like watching a World Series ball game from a seat behind a stadium post. The comet fragments would nevertheless strike Jupiter close to the side facing Earth, so astronomers hoped that something would be seen when the planet's rapid rotation, of once every 9 hours 55.5 minutes, brought the impact sites into view. Moreover, the *Galileo* spacecraft, on its way to Jupiter, had a direct view of the actual collisions from its unique position in space.

No one was disappointed! When the comet fragments plowed into Jupiter, each of them exploded with an energy equivalent to the simultaneous explosion of hundreds of thousands of nuclear bombs on Earth. As it penetrated the outer atmosphere of Jupiter, each comet fragment heated and compressed the surrounding gas, producing a violent explosion high in the atmosphere. The resultant fireball punched a hole through the overlying material and sent plumes of hot gas rising into space. Using instruments on the *Galileo* spacecraft, they were detected at the side of Jupiter by their infrared heat radiation, indicating temperatures of up to 20 000 kelvin. It took 10 or 20 minutes for each plume to rise and fall again, by which time the impact site had rotated into view from Earth.

The comet fragments slammed into Jupiter, one after another, like the cars of a train when its locomotive is derailed. After generating a bright ball of light, each fragment disfigured Jupiter with a black scar that had never been seen before, twice as large as the Earth and spanning tens of thousands of kilometers (Fig. 13.3). Meanwhile, waves swept across the impact site and reverberated deep within the planet, which seemed to shudder from the impacts. The dark marks endured for months, gradually spreading, merging and slowly fading from view as the winds in Jupiter's atmosphere dispersed their material.

13.2 Consumed by the Sun

Some comets plunge deep into the Sun's thin million-degree outer atmosphere, or corona. Instruments aboard the *SOlar and Heliospheric Observatory* (*SOHO*) satellite have recorded their death-defying trip around the Sun. One of its instruments uses an occulting disk to block out the bright light of the visible solar disk, enabling it to detect the comets as they move through the inner corona (Fig. 13.4). A comet often pays a heavy price for this trip, sometimes breaking apart because of the Sun's gravitational forces.

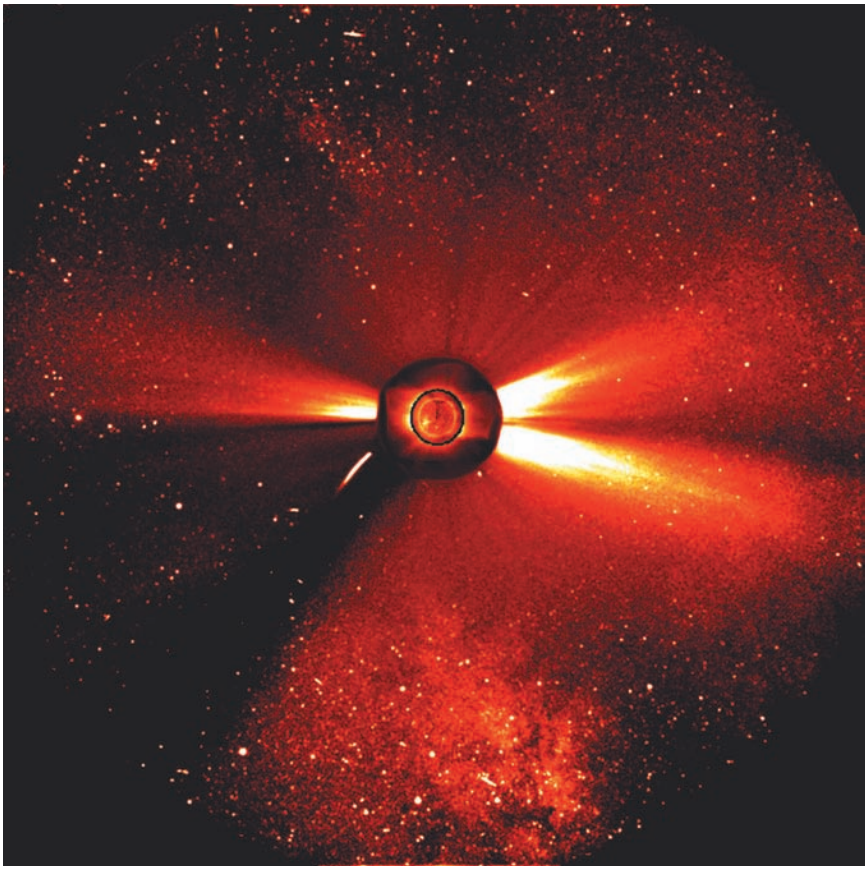


Fig. 13.4 Fatal impact of a comet into the Sun This composite image records a comet plunging into the Sun on 23 December 1996. The innermost image (*center*) records the bottom of the million-degree solar atmosphere, known as the corona. The electrically charged coronal gas is seen blowing away from the Sun just outside the inner dark circle, which marks the edge of one instrumental occulting disk. Another instrument records the comet (*lower left*), as well as the coronal streamers at more distant regions and the stars of the Milky Way. (Courtesy of the *SOHO* EIT, UVCS and LASCO consortia. *SOHO* is a project of international collaboration between ESA and NASA.)

Other comets are hurtling toward complete meltdown, passing so close to the Sun that the encounter is fatal. Though rarely, if ever, hitting the visible solar disk, or photosphere, these comets can come closer than 50 000 kilometers from it. They are unlikely to survive the Sun's intense heat and gravitational forces at that range.

Most of the comets discovered by *SOHO*, about 90 percent of them, are small comet fragments known as the Kreutz sungrazers, which closely approach the Sun from one direction in space. They are named after the German astronomer Heinrich Kreutz (1854–1907) who found that many of the comets that traveled exceptionally near to the Sun in the 19th century seemed to have a common origin with similar orbits. They are all probably fragments of a single large comet that first broke up when passing very close to the Sun thousands of years ago.

When a member of the Kreutz sungrazer group moves around its orbit and returns to our vicinity, it can dive into the inner corona and disappear forever (Fig. 13.5). Each comet fragment can be very small, just 6 to 12 meters across, despite their spectacular display. Such a tiny object, falling so close to the Sun, would vaporize completely away.

The occulting instrument aboard *SOHO* has unexpectedly watched more than a thousand comets pass near the

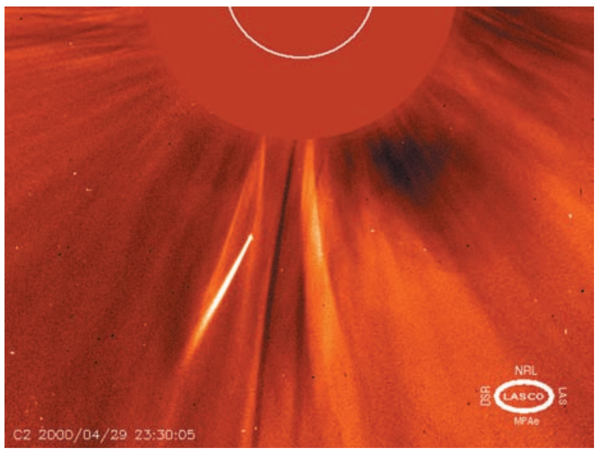


Fig. 13.5 Death of a sungrazer This *SOHO* image shows a bright sungrazing comet (*bottom center*) headed into the inner atmosphere of the Sun on 22 October 2001. The partial white circle marks the outer edge of the visible solar disk, whose intense glare is hidden by the instrument's occulting disk (*opaque circular region at top center*). The million-degree solar atmosphere can also be seen, steaming away from the Sun. (Courtesy of the *SOHO* LASCO consortium. *SOHO* is a project of international collaboration between ESA and NASA.)

Sun or even into it. Amateur astronomers from all over the world have examined *SOHO*'s real-time images, posted on the Internet at http://sohowww.nascom.nasa.gov/, discovering hundreds of previously unknown comets, including numerous Kreutz sungrazers and many other new comets as well.

The collisions of comets with Jupiter and the Sun have helped raise worldwide awareness of a similar threat to our home planet. Such impacts have happened on Earth in the past and they could happen again, with devastating effects to civilization.

13.3 Impacts of asteroids with the Earth

The solid surfaces of the terrestrial planets, as well as the Earth's Moon, are marked with impact craters, the scars of past collisions with cosmic objects speeding through space. They originated by the coalescence of these objects, and the barrage continues at a lower rate today when many cosmic objects travel across the orbits of the terrestrial planets (Fig. 13.6). So the Earth does not occupy a secure niche in space. Our planet is instead immersed in a cosmic shooting gallery, subject to a steady bombardment by potentially lethal, Earth-approaching objects.

The large majority of asteroids are in orbits between those of Mars and Jupiter where they pose no threat to Earth. Some of these rocky objects, however, follow a more eccentric course that takes them closer to Earth, sometimes crossing its path. And although most short-period comets do not come closer to the Earth than the nearest planets, some of these icy intruders can pass perilously near to us. It is these threatening asteroids or comets, collectively known as Near-Earth Objects (NEOs), which may be on a collision course with the Earth. If one of these cosmic bombs hits our planet, it could explode with a violence that far surpasses the world's entire nuclear arsenal, threatening civilization and possibly making humans extinct.

Explosions in the atmosphere

We have dramatic evidence that intruders from space are constantly bombarding the Earth today. Unknown to the public, the US Department of Defense has been detecting them for decades. Military satellites, designed to watch for enemy rocket firings and nuclear explosions, have detected the explosions produced when speeding, house-sized cosmic objects enter the air. The incoming projectiles are heated to incandescence and then self-destruct in the upper atmosphere, vanishing without a trace on the ground.

Ground-based defense networks, designed to listen to the sounds generated by man-made nuclear explosions, have confirmed the satellite results. They show that one of these cosmic bombs is bursting overhead every month, each with an energy equivalent to a nuclear bomb. These investigations might be aiding world peace by helping to distinguish between natural explosions and those caused by humans, thereby preventing false warnings of clandestine nuclear tests or of nuclear attack by terrorists.

The largest object to strike the Earth in the 20th century wasn't quite big enough to reach the ground. It disintegrated between 5 and 10 kilometers up, over the Podkamennaya Tunguska River in central Siberia. The shock wave generated by the ensuing explosion leveled trees over 2 trillion (2×10^{12}) square meters of the underlying land, an area larger than New York City and surrounding suburbs (Fig. 13.7). The energy produced was equivalent to the aerial explosion of the nuclear bomb that leveled Hiroshima. So much devastation, yet it failed to produce a crater.

A Tunguska-like atmospheric explosion, relatively near the ground, is the usual fate for a stony asteroid fragment roughly 50 meters across. An asteroid of this size would have had the internal strength to penetrate deeply into the Earth's atmosphere before exploding, but a comet of comparable size would have disrupted too high in the atmosphere to cause much damage to the ground.

Scientists estimate that a Tunguska-like explosion should occur every century or two, on average. If it happens above a large city, the results will be disastrous. But then, it is more likely to enter the atmosphere above the oceans or remote land regions, since they cover much more of the planet's surface area. Terrestrial impacts by larger cosmic objects occur less frequently, but they will reach the ground with more powerful consequences.

The terrestrial impact record

Cosmic intruders just tens of meters in size can survive passage through the Earth's atmosphere more or less intact and strike the terrestrial surface at high velocity. Even a relatively small metallic asteroid, just tens of meters across, is tough enough to penetrate the atmosphere and hit the ground. Stony asteroids have to be more than 50 meters in size to survive passage through the atmosphere, and the fragile comets must be at least ten times as big to make it through.

A classic example is the 1.2-kilometer-wide Meteor Crater in northern Arizona, also known as the Barringer Crater (Fig. 13.8). It was formed about 50 000 years ago by the impact of a nickel-iron asteroid just 60 meters in diameter. When it struck the Earth, this relatively small projectile released an estimated 100 million billion (10^{17}) joules of kinetic energy, an amount comparable to the

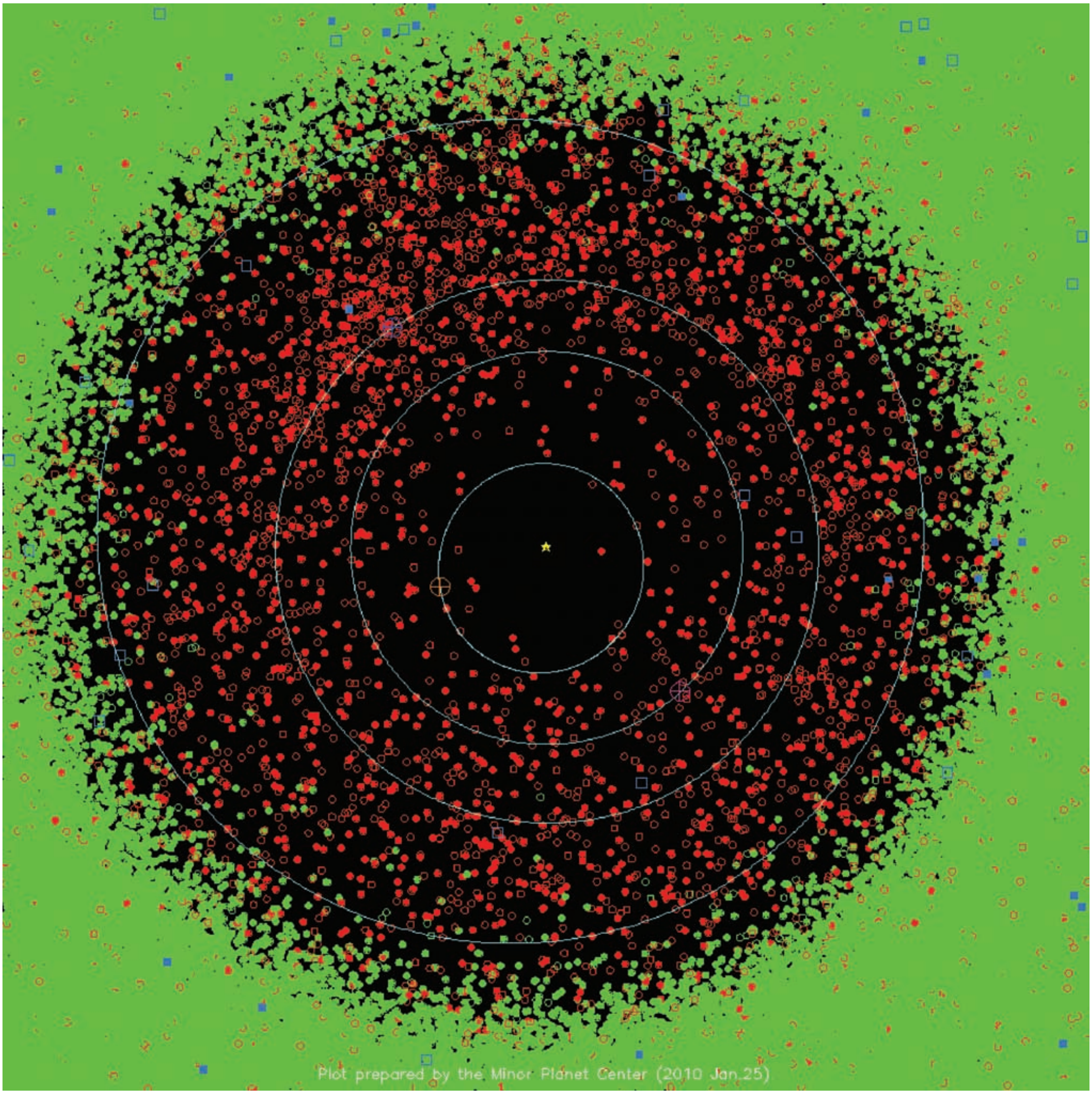


Fig. 13.6 Minor and major planets in the innermost solar system The orbits of the inner major planets are shown as large light-blue circles, for Mercury, Venus, Earth and Mars from the center out. Small red circles indicate the locations of the minor planets coming within 1.3 AU of the Sun at perihelion, where 1.0 AU is the mean distance between the Earth and the Sun. Small filled circles indicate objects observed at more than one opposition. Outline circles indicate objects seen at only one opposition. Numbered periodic comets are shown as filled light-blue squares. Other comets are shown as unfilled light-blue squares. The plot is for the date 20 January 2010. (Courtesy of Gareth Williams, Minor Planet Center.)

energy released by the explosion of 20 million tons of TNT (trinitrotoluene).

Even the largest craters, produced by the biggest asteroids, will gradually disappear from sight with the passage of time. The same forces that erode mountains, deposit sediments, eject lava and shift continents are erasing the craters and removing them from sight. If not for these dynamic forces, the craters accumulated over the ages on the Earth would be as densely distributed and prominent as the overlapping craters on our Moon.

Only about 176 terrestrial impact craters have managed to survive the ravages of time. They have been identified

Fig. 13.7 The Tunguska explosion On 30 June 1908, a giant blue-white ball of fire streaked across the daytime sky above the Tunguska River in Siberia, apparently becoming brighter than the Sun. The Tunguska fireball then exploded in the atmosphere, felling underlying trees like matchsticks. All of the toppled trees point away from a central location that must have been directly below the point of the explosion. Interpreting the extent and orientation of the tree-fall pattern, scientists concluded that the explosion above Tunguska released an energy equivalent to exploding a nuclear bomb between 5 and 10 kilometers above the Earth's surface.

on images taken from space, using airplanes, the *Space Shuttle*, or satellites such as *Landsat*. These craters can be first identified from aerial photographs, by their circular shapes and uplifted and overturned rims (Fig. 13.9). Aerial images of about one quarter of the identified Earth impact craters have been provided by scientists at the Lunar and Planetary Institute on the Internet at http://www.lpi.usra.edu/publications/slidesets/craters/.

But since other processes, such as volcanism and erosion, can also leave circular holes, confirming evidence of an impact origin must be gathered from rocks in and around the crater. How do geologists know that some terrestrial craters are due to the explosions of projectiles coming from space? They look for rocks that have been transformed under the conditions of extreme temperature, pressure, and shock associated with a high-velocity external impact. The most apparent shock effect is the formation of conical structures called shatter cones, which point toward the center of the impact (Fig. 13.10). Other evidence includes glassy, previously molten material formed at high temperature, and minerals with a deformed crystal structure produced by a shattering, high-pressure impact. Roughly 25 percent of the craters also contain meteorites that had to come from space.

Geologists have dated some of the craters by radiometric age determinations of previously melted rock, determining the time that has elapsed since the molten rock cooled and solidified. They find ages from a few thousand to 2 billion years, but most of them are younger than 200 million years old. The lunar crater record suggests that the cratering rate on Earth must have been roughly constant during the past 3 billion years, so erosion and other geological processes have worn most of the older ones away.

Only the largest terrestrial craters have been able to survive the wearing effects of time for more than a few million years. One of the largest and oldest ones, for example, is located near Vredefort, South Africa; it is 300 kilometers in diameter and it was formed about 2 billion years ago.

Fig. 13.8 Meteor crater One of the best and earliest known impact craters on Earth, located in the northern Arizona desert. It has been named the "Meteor Crater" for the nearby post office named Meteor, but it is also widely known as the Barringer Crater in honor of Daniel Barringer (1860–1929), a mining engineer and business man, who was the first to suggest, in 1903, that it was produced by meteorite impact. The crater, about 1.2 kilometers in diameter, was excavated about 50 000 years ago during the impact of a nickel-iron meteorite about 50 meters across. When first discovered, about 30 tons of meteoritic iron was found scattered about the crater. The meteorite is officially named the Canyon Diablo Meteorite, after the nearby ghost town of Diablo Canyon. (Courtesy of NASA.)

Fig. 13.9 Wolf Creek impact crater This relatively well-preserved crater near Wolf Creek, Australia, is partly buried under windblown sand. Iron meteorites have been found in the vicinity, as well as some impact glass. The rim diameter of this crater is 850 meters and the impact that created it occurred about 300 000 years ago. (Courtesy of Virgil L. Sharpton and the Lunar and Planetary Institute.)

Table 13.1 The 10 largest identified terrestrial impact craters[a]

Crater name	Location	Diameter (km)	Age (My)	Meteoritic component
Vredefort	South Africa	300	2023 ± 4	Chondrite
Sudbury	Ontario, Canada	250	1850 ± 3	–
Chicxulub	Yucatán, Mexico	170	64.98 ± 0.05	Chondrite
Popigai	Russia	100	35.7 ± 0.2	Chondrite
Manicouagan	Quebec, Canada	100	214 ± 1	–
Acraman	South Australia	90	~ 590	Chondrite
Chesapeake Bay	Virginia, USA	90	35.3 ± 0.1	
Puchezh-Katunki	Russia	80	167 ± 3	
Morokweng	South Africa	70	145.0 ± 0.8	Chondrite
Kara	Russia	65	70.3 ± 2.2	Chondrite?

[a] The crater diameters are in units of kilometers (km), and the ages are in millions of years (My).

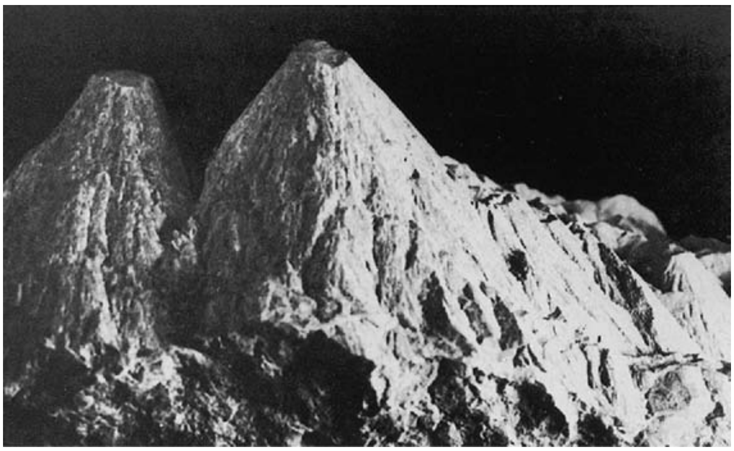


Fig. 13.10 Shatter cones The shatter cones that are found in the vicinity of many terrestrial craters provide evidence for shocks associated with the impact of a cosmic object. They point towards the direction of impact, like the cone-shaped plugs of glass that are often formed when a bullet strikes a window. The shatter cones shown here are about 0.05 meters in height. They are from the Wells Creek Tennessee Basin, a crater that is about 14 kilometers across and roughly 200 million years old.

An extensive database of known terrestrial impact craters is maintained by the Planetary and Space Science Center, University of New Brunswick, Canada, at http://www.unb.ca/passc/ImpactDatabase, listing them by location, age, diameter or name with other supplemental information including images. Many of the 10 largest ones have been identified with chondrite meteorites and they are between 35 million and 2 billion years old (Table 13.1).

The consequences of these bigger impacts are even more sobering than the small ones. If an asteroid of just 1 kilometer in size hit the Earth, the power of its explosion could not be matched by the world's entire nuclear arsenal. Such an impact is estimated to occur every one million years or so. An asteroid of exceptional size, say 20 kilometers across, might hit the Earth less often, every 100 million years, on average. As we shall next see, such collisions have happened in the past, when they altered the course of biological history.

13.4 Demise of the dinosaurs

Catastrophe from the sky

Collisions by objects from outer space have always been a menace to life on Earth. During the planet's first billion years, the barrage was probably so intense that living things could not exist on the Earth's surface. After those early times, the rate of bombardment slowed down, so impacts of exceptionally large cosmic projectiles became less frequent. But these giant impacts continued every once in a while, with devastating consequences. The most recent death-rock arrived 65 million years ago, wiping out the dinosaurs and many other living things. Such an abrupt destruction of an entire species by a force of nature is known as a mass extinction.

A thin worldwide layer of iridium-rich clay, just 0.01 meters thick, provided the initial evidence that a cosmic collision wiped out the dinosaurs. When a team headed by the American geologist Walter Alvarez (1940–) determined the layer's age, from its position among geologically dated strata, they found that it was deposited about 65 million years ago when the dinosaurs and a variety of plants and animals disappeared.

The clay layer contained unusual amounts of the rare element iridium. It is not found in such quantities in terrestrial rocks since most of the primordial iridium on Earth sank to the planet's core during differentiation early in its history. In contrast, iridium is much more abundant in certain meteorites, so they might have supplied the clay layer with iridium.

Since the cosmic iridium rains steadily down through the atmosphere and settles in the soil, the amount of iridium in a layer of sediment can be used as a cosmic clock. In an average century, a certain amount of iridium will mix with the soil and become part of any new layers that are forming. If a layer requires twice as long to form, it will have twice as much iridium.

But the geologists found that the iridium clock had gone wild for a short interval about 65 million years ago. The amount of iridium they found in this layer of clay was far higher than normally found in the Earth's crust and about 30 times higher than that found in the fossilized limestone above and below the clay. Moreover, the same type of iridium enrichment was found in clay at widely scattered points on the Earth. So the entire globe had been drenched with an unusually large amount of iridium for a short time.

Walter, his father Luis Alvarez (1911–1988), and their colleagues concluded in 1980 that the iridium deluge came from outside the Earth, delivered by a large asteroid that struck the Earth and vaporized about 65 million years ago. According to their hypothesis, the iridium was lofted into the atmosphere along with other debris by the fireball of hot gas created during the collision, and then carried by the winds over much of the globe. The worldwide cloud of iridium-rich dust then slowly filtered back down to the ground where it produced a thin global layer that contained relatively large amounts of iridium. They estimated that a layer 0.01 meters thick covering the entire Earth would be deposited by an asteroid about 10 kilometers in diameter.

Most geologists and biologists must have initially dismissed the idea of a killer asteroid from space. They probably attributed it to spaced-out astronomers or science-fiction enthusiasts. Such an abrupt cataclysm conflicted with the prevalent concept of gradual evolutionary change over the eons. Sudden, short-lived events were just not supposed to affect the course of evolution. Yet the evidence for catastrophic events is found throughout the solar system, including the cratered surfaces of the Earth and its Moon.

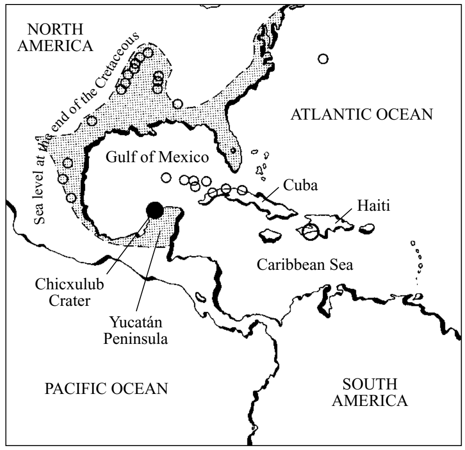


Fig. 13.11 Site of impact that wiped out the dinosaurs The long-sought site of the scar left by a killer asteroid has been found near Chicxulub (*filled circle*), a small village at the tip of the Yucatán Peninsula. Material from the submerged crater has been dated at about 65 million years ago, coinciding with a blast that triggered the eradication of most life on Earth. Thick sedimentary deposits laid down 65 million years ago in Haiti (*big open circle*) contain exceptionally large amounts of iridium, shocked quartz and glassy debris that are thought to be part of the impact across the Caribbean, almost 2000 kilometers away. The small open circles mark the sites of marine wave deposits associated with the same impact. [Adapted from Alan R. Hildebrand and William V. Boynton, *Natural History* **6** 47–53 (1991).]

The idea just would not go away and supporting evidence kept accumulating. Shocked quartz grains were found in the iridium-rich clay layer worldwide. No known terrestrial process, including volcanic flows or explosions, can generate pressure high enough to alter the grains in the observed way – only the sudden shock of an impact can. Still, the skeptics asked: Where is the crater of the impact that occurred 65 million years ago and was big enough to obliterate most of the Earth's life forms?

After years of searching, the telltale crater was found straddling the northern coastline of the Yucatán Peninsula (Fig. 13.11). It is located below the Mayan village of Chicxulub (pronounced Cheek-shoe-lube, a Mayan phrase for "horns of the devil"), and is hence known as the Chicxulub impact basin. The discovery of this crater and the subsequent confirmation of its age at 65 million years led to the widespread acceptance of the impact hypothesis for the demise of the dinosaurs.

At the time of the impact, that part of the Yucatán Peninsula was below sea level, and the center of the crater now lies buried below 1.1 kilometers of limestone laid down in the intervening years. So there is nothing on the surface to betray the crater's existence, and the vast scar cannot be seen directly.

Focus 13.1 The belt of an asteroid

The energy of the collision of an asteroid with the Earth can be calculated from the kinetic energy (K.E.) of the impacting object, which is given by the expression

$$\text{K.E.} = \frac{1}{2} M V^2$$

where M is the mass of the projectile and V is its incoming velocity. The mass of an impacting asteroid can be determined from the mass density and volume. Assuming a mass density, ρ, of about $\rho = 3000$ kilograms per cubic meter, comparable to that of stony asteroids and meteorites, and a radius, R, of $R = 5$ kilometers, a mass of

$$M = \rho \times \frac{4}{3}\pi\ R^3 = 1.6 \times 10^{15} \text{ kilograms}$$

is obtained. Most near-Earth objects travel with orbital velocities of about 30 kilometers per second, comparable to the Earth's, and using this velocity with the mass we have estimated, we obtain

$$\text{K.E.} = 7.2 \times 10^{23} \text{ joules}$$

This amount of energy is equivalent to the explosion of nearly 100 000 terrestrial bombs, each with a destructive force of 100 megatons (100 million tons or 10^{11} kilograms) of trinitrotoluene (TNT). The destruction produced by the impact would therefore be many orders of magnitude larger than that caused by the simultaneous explosion of the world's entire arsenal of nuclear weapons. It would be equivalent to the detonation of the blast that destroyed Hiroshima, at just 13 000 tons of TNT, every second for 175 years.

The serendipitous discovery of the submerged crater began with a search for oil in the region. The Mexican national petroleum company, Petróles Mexicano or Pemex for short, commissioned an aerial magnetic survey to assess the thickness of sedimentary – and possible oil-bearing – rocks in the region, which revealed a large, buried semicircular structure. Coarse gravity maps showed a similar feature, and exploratory oil-drilling in the area revealed an underground layer of broken, melted rock. Subsequent radioactive dating of the drill-core fragments showed that they were exactly contemporaneous with the clay layer, with an age of 65 million years.

Chicxulub is the biggest thing to hit Earth in the past 1 or 2 billion years, and of just the right size to have been excavated by an asteroid of 10 kilometers in extent. The energy released during the collision of such an intruder is enough to trigger a mass extinction (Focus 13.1).

The day the dinosaurs died

Most scientists are now convinced that the dinosaurs, which had dominated the Earth for over 160 million years, were destroyed when a marauding asteroid dropped out of the sky and struck the Earth 65 million years ago. The explosive impact generated an enormous ball of fire, which incinerated everything in the immediate area, blasted billions of tons of debris into space, and gouged out a crater 170 kilometers across. But that was just the beginning.

The cosmic intruder was completely vaporized during the explosive impact, and a great fireball rose into the stratosphere, carrying with it large amounts of pulverized debris. Vast clouds of dust and ash remained suspended and were circulated by air currents until they encircled the Earth, covering it in total darkness. Since the global shroud blocked the Sun's heat, as well as its light, the surface temperature plummeted and the planet entered a dark chill, lasting for months before the dust eventually settled back to Earth.

Plants would have failed to receive enough sunlight to allow photosynthesis to continue, and a prolonged "winter" of unusual cold would have added to the devastation. As plant life withered and froze, plant-eating animals dependent on them would also die, as would meat-eating animals once their plant-eating prey were gone. By the time that the dust settled, many land plants and animals were no longer there.

The cosmic object that slammed into the water above the current Yucatán Peninsula would have also generated enormous sea waves, great moving walls of water that surged into the East Coast of what is now known as the United States, leveling everything in their path and leaving almost nothing standing.

Other devastating effects most likely included: extensive wildfires, ignited when hot material ejected from the impact fell back down to the ground; acid rain that poisoned the water and destroyed shell-bearing creatures, disrupting the entire food chain; and volcanoes that might have contributed to the mass extinction by spewing carbon dioxide, deadly ash, sulfur and other substances that disrupted the atmosphere and altered the climate. When combined with the dust, sulfur, and hot debris tossed skyward by the original impact, the volcanic activity could have spelled doom for many living things.

When the cosmic blast and its aftermath were over, the dinosaurs were gone, along with most marine animals and many land plants. Altogether roughly half of all animal and plant species were wiped out. That sounds pretty awful, but from catastrophes there arise opportunities. The biological devastation caused by the impact apparently cleared the way for the rise of the relatively small mammals, so your

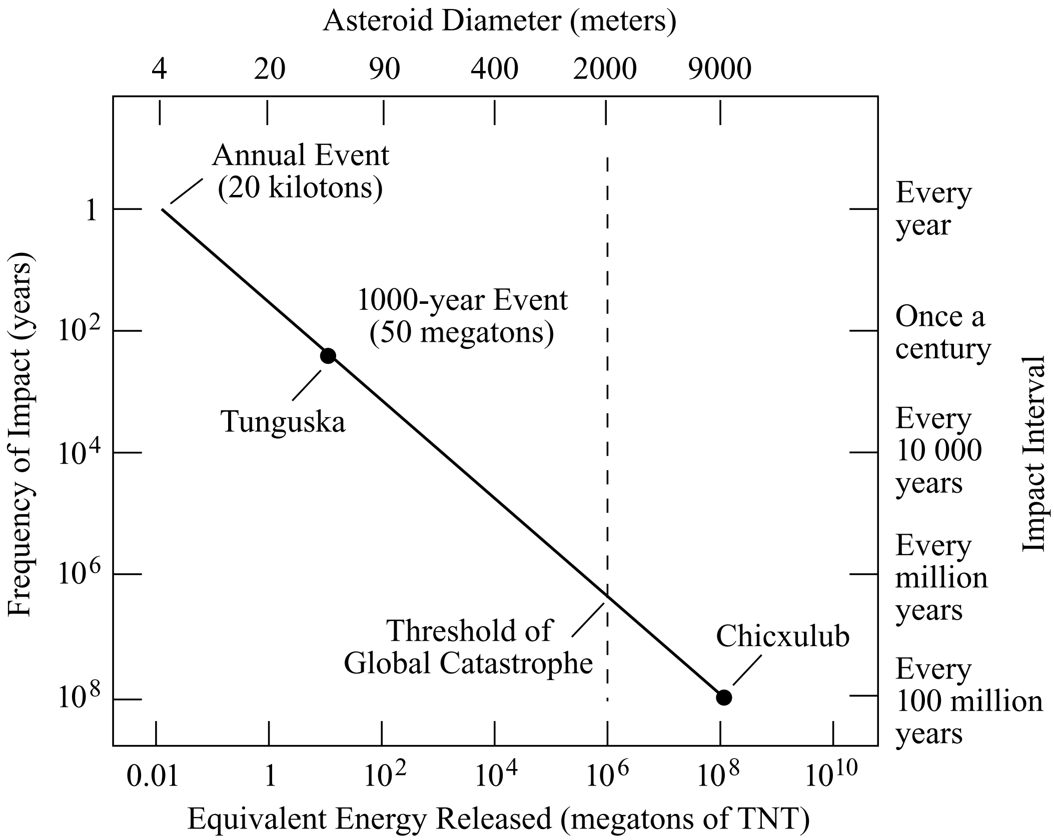


Fig. 13.12 Cosmic impact probabilities A hail of cosmic objects continually pelts the Earth from space. This plot shows the likelihood (*vertical axis*) that a member of the current population of Earth-crossing asteroids will hit our planet. An object two kilometers in diameter, capable of producing certain worldwide damage, hits the Earth every million years on average. An impact like the one that wiped out the dinosaurs, giving rise to the Chicxulub crater, is estimated to occur every 100 million years. Since smaller asteroids are much more numerous than larger ones, the smaller objects strike our planet more often. An impact like the Tunguska event, which occurred on 30 June 1908, might occur every 1000 years or so. The explosive energy of the impact is also given (*horizontal axis*) in units of megatons of exploded TNT. One hundred megatons (10^{11} kilograms) of TNT is equivalent to about 4×10^{17} joules of energy, the amount released by a typical nuclear bomb on Earth.

very distant ancestors may have benefited. Some scientists have even argued that killer asteroids may periodically sweep a wave of death across the Earth, thereby ending the rule of the dominant species.

We humans have flourished in the past half million years, developing wonderful civilizations, building great cities, generating profound knowledge, and sending spacecraft throughout the solar system. Yet it might suddenly come to an end. Such catastrophes have a small probability of occurring during our relatively short lives, but over astronomical times of millions and billions of years, the exceedingly unlikely becomes a virtual certainty. After all, 99 percent of all species that ever lived have gone extinct.

13.5 Assessing the risk of death from above

Somewhere in space, an asteroid is hurtling toward a future collision with Earth. And if it is large enough, the impact will severely disrupt terrestrial life. It's only a matter of time.

So when are we going to be hit hard enough to worry about it? Since astronomers have not yet located the doomsday rock that will definitely collide with Earth, we don't know exactly when the next impact will take place. But we can calculate the odds, and they depend on the size of the colliding object. Since there are many more small cosmic objects than large ones, the smaller ones hit our planet more frequently. Bigger asteroids strike the Earth less often, but they can cause vastly greater damage.

Thus, to estimate the risk of being hit in a way that matters, the potential impacting projectiles first have to be sorted according to size (Fig. 13.12). Fragments smaller than a few tens of meters across burn up in the atmosphere and rarely reach the ground. Asteroids a hundred meters in diameter are expected to strike Earth every thousand years on average. They could take out a city and cause

Table 13.2 The dangers of a lifetime[a]

Cause of death	Chance of dying in a 65-year period	Number of people killed
Car accident	1 in 100	5
Murder	1 in 300	1
Gun accident	1 in 2500	1
Accidental electrocution	1 in 5000	1
Asteroid impact (two kilometers in diameter, global catastrophe)	**1 in 20 000**	**100 million**
Airplane crash	1 in 20 000	300
Tornado	1 in 60 000	10
Snake bite or bee sting	1 in 100 000	1
Food poisoning	1 in 3 000 000	1

[a] Adapted from Clark R. Chapman and David Morrison, *Nature* **367**, 33–40 (1994). The risks are for a person living in the United States.

severe local damage, but pose no threat to the Earth as a whole.

Asteroids about one kilometer in size pose a greater peril. They are large enough to destroy a large country and produce global consequences. Contemporary surveys indicate that there may be about one thousand of these objects now on paths that come near the Earth's orbit. Most of the time, the Earth will be somewhere else if one of them crosses its path, but occasionally they will arrive almost simultaneously at the intersection. The average time between such impacts is about one million years, an interval vastly longer than the history of civilization.

During your lifetime, the chance that you will be wiped out by the impact of a two-kilometer asteroid is 1 in 20 000, the same as death from an airplane crash (Table 13.2). But it is not just one person that dies when a large cosmic projectile hits the Earth. About 100 million people are expected to perish if the object is two kilometers across, and a bigger one could destroy all of us. The other causes of death usually affect just one person at a time, or perhaps a few of them.

Moreover, to declare that cosmic disaster strikes the Earth, on average, just once in a million years or so does not mean that we are guaranteed such a long interval between catastrophic impacts. The very small chance of such a collision is the same today as it will be millions of years from now. It could be tomorrow, or it might be long after you're gone.

In fact, there are all sorts of other things you can worry about if you are in a morbid mood. The chance of dying before age five is 1 in 8 in South Asia, and 1 in 100 in the United States. The risk of dying in childbirth in the United States is about 1 in 12 000, roughly twice the chance of death by the impact of a two-kilometer cosmic projectile, but thousands of times greater than death by dog bite or by drinking detergent.

But to get back to the dangers of cosmic impact, it takes an even larger, rarer asteroid, 10 kilometers across or larger, to render the human species extinct. The impacting object would blast out a huge crater and eject billions of tons of pulverized rock and dust into the air. The globe-circling pall of dust and other debris would block out the Sun's light and heat, crippling agriculture, producing widespread starvation, and perhaps leading to a worldwide breakdown of our fragile civilization.

Such a mass extinction might occur once in a hundred million years on average, destroying in an instant what it has taken humans millennia to build. The chances of that happening are very low, but they are not zero. The lifetime risk of your being wiped out with the rest of humanity during such a mass extinction is roughly one in a million. So it might happen, and a prudent society should prepare for the possibility.

13.6 Breaking a date with doomsday

Finding the hidden threat

While we know that asteroids and comets have collided with the Earth in the past, and that they will inevitably hit our planet in the future, we do not yet know if any them are now headed for a deathly collision with our solitary outpost of life. Astronomers are therefore taking a census of everything out there that is big enough and close

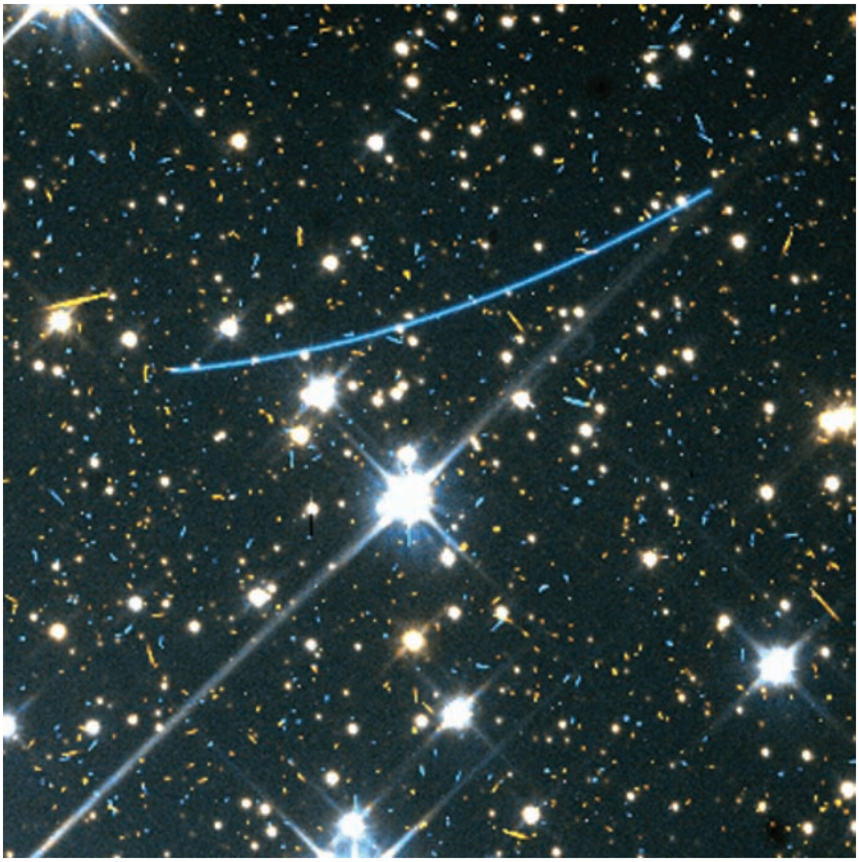


Fig. 13.13 Asteroid streak As an asteroid moves along its orbit it will produce an elongated trail in an image taken with a telescope following the background of "fixed" stars, which are not moving in this way. One previously unknown asteroid was discovered in 1998 as a long blue streak (*top center*) against the white background stars in this archival image taken by the *Hubble Space Telescope*. The bright asteroid has an estimated diameter of 2 kilometers, and was located at about 140 million kilometers from the Earth. (Courtesy of NASA/R. Evans and K. Stapelfeldt, JPL.)

enough to threaten us (Figs. 13.13, 13.14, 13.15). Once all of these Near-Earth Objects (NEOs) are located, and their current trajectories known, astronomers can use computers and refined observations to determine their precise future paths and establish whether and when any of them will strike the Earth.

After decades of scanning the skies, ongoing search programs will find and catalog the most threatening asteroids (Focus 13.2). They will provide the exact positions and orbits of about 90 percent of the near-Earth asteroids larger than one kilometer in size. That is the minimum diameter of a space rock that could have global consequences if it hit Earth.

Astronomers have found nearly one thousand near-Earth asteroids with diameters of one kilometer or larger, and they seem to account for the majority of the largest space rocks that might wreak global havoc. None of those discovered so far is on a direct collision course with Earth – at least in the near future. The search is therefore slowly diminishing the chances of our demise, gradually improving the odds of this cosmic Russian roulette.

Just as there are more small fish than large ones in the sea, there are many more small cosmic objects that are now moving about in the space near our planet, and there are also about one thousand of them that are already known to be potentially hazardous. Most of these probably pose

Fig. 13.14 Near-Earth asteroid The red dot at the center of this image is the first near-Earth asteroid discovered, on 12 January 2010, with the *Wide-Field Infrared Survey* (*WISE*) spacecraft. The asteroid is about 1 kilometer in diameter, and moves on an elliptical orbit that is inclined to the plane of our solar system. This orbit takes the asteroid beyond Mars and as close to the Sun as the Earth. The all-sky *WISE* infrared survey, which began on 14 January 2010, is expected to find about 100 000 previously undiscovered asteroids in the main belt between Mars and Jupiter, and hundreds of near-Earth asteroids. It will reveal the darkest asteroids, which don't reflect enough visible sunlight to be seen with conventional optical telescopes, but heat up and emit the infrared radiation detected from *WISE*. (Courtesy of NASA/JPL-Caltech/UCLA.)

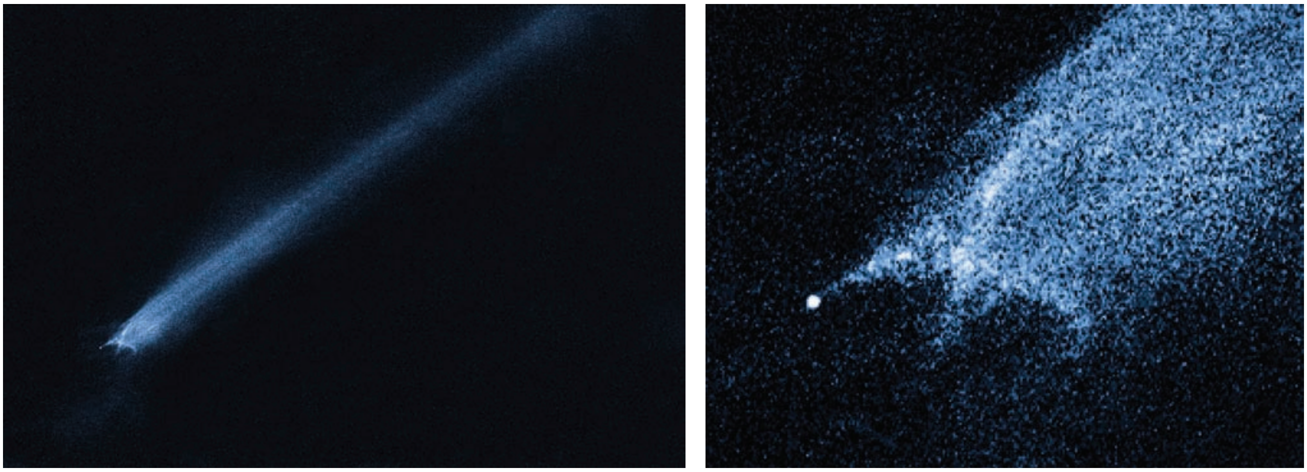


Fig. 13.15 Suspected asteroid collision leaves trailing debris This unusual object was first discovered by the ground-based LIncoln Near Earth AsteRoid (LINEAR) program, designed to detect asteroids that may collide with the Earth in the future. When the *Hubble Space Telescope* zoomed in to take a look with its sharper vision, it found a comet-like tail of debris (*left*), flowing from an X-shape seen in the close-up view (*right*) of the front center of the tail. The X may mark the spot of an asteroid collision in the main belt of asteroids that created the debris. The remains of one of the asteroids, a nucleus 140-meters in size, are offset from the tail center and possible collision site. (Courtesy of NASA/ESA/D. Jewitt, UCLA.)

a significant threat, but not a global one. They can hit the Earth with an energy equivalent to a large nuclear bomb.

Today it is beyond our technology to detect and defend against the smaller cosmic bombs of several meters across, but they can be seen when coming very near the Earth. Some of these asteroids have come uncomfortably close, passing within less than the mean distance between the Earth and its Moon, which is 384 400 kilometers away. By cosmic standards, that is a close call. About 100 asteroids have passed this close to the Earth in the past several years, most of them less than 200 meters across.

On 13 January 2010, for example, an asteroid of 10 to 20 meters in diameter passed within one-third the distance from the Earth to the Moon. Others have nearly hit our planet, passing within a few Earth radii (Table 13.3), sometimes with just a few hours notice or with no advance warning at all.

Table 13.3 Very close approaches of asteroids to Earth[a]

Closest distance from Earth (R_E)	Date of closest approach	Estimated diameter (meters)
1.0	7 October 2008	2 to 5
2.0	31 March 2004	5 to 12
2.1	9 October 2008	0.5 to 1
3.2	6 November 2009	5 to 12
4.8	20 October 2008	5 to 12
5.3	19 December 2004	1 to 3
7.1	3 November 2008	2 to 5
7.6	8 March 2004	20 to 50

[a] The radius of the Earth is $R_E = 6378$ kilometers.

Doing something about the threat

The identification and deflection of a near-Earth asteroid that is heading toward a direct collision with Earth is now regarded as a global public safety issue. By 2010 NASA had identified more that 1000 potentially hazardous NEOs, but a full inventory was still incomplete, and the White House has urged the development of evasive action to deflect one when we see it coming, perhaps related to the goal of landing an astronaut on an asteroid by 2025.

How much warning will we have? When the inventory of large NEOs is complete, and one is found on its way to strike the Earth, then warning would probably come decades in advance of a collision. The threatening asteroid would most likely swing near the Earth and loop around the Sun several times before hitting our planet. Since existing surveys do not regularly detect small NEOs, less than 200 meters in size, as efficiently as they discover larger ones, we may not know about them until they are about to collide with us. Thus we will either have a long lead-time or none at all.

And what do we do if we find a large cosmic projectile headed our way? The Earth cannot be moved out of the way, but we could launch an intercept mission to redirect

Focus 13.2 Searching for cosmic bombs headed our way

The collision of a comet with Jupiter, anticipated in 1993 and watched by millions in 1994, raised public consciousness of the impact threat to planet Earth. The US Congress held hearings to study the threat, and asked NASA to formulate plans to deal with the problem. Public awareness of the cosmic bombs was notched up once more with the release, in 1998, of two blockbuster movies, *Deep Impact* which deals with a killer comet and *Armageddon* in which two astronauts save humanity by diverting a rogue asteroid headed toward collision with Earth. In the same year, the US Congress held more hearings about the threat, and NASA initiated the Spaceguard Survey, intended to find, within the next decade, 90 percent of the Near-Earth Objects (NEOs) which might be on a collision course with Earth and are larger than one kilometer.

By the end of the 20th century, several teams of astronomers were surveying the sky with electronic detection equipment and computers to complete the inventory of large NEOs, supported by NASA and the US Air Force. The search involves finding every object that moves against the background stars down to 19.5th magnitude, almost 100 000 times fainter than the detection limit of the human eye.

Exceptionally productive search programs, resulting in the largest number of discoveries of near-Earth asteroids, are the LIncoln Near Earth Asteroid Research (LINEAR) project of MIT's Lincoln Laboratory and the Catalina Sky Survey, based at the University of Arizona. They have together discovered between 100 and 400 near-Earth asteroids every year between 1998 and 2010, and nearly one thousand of these are thought to have diameters of one kilometer or larger.

More recently, the *Wide-field Infrared Survey Explorer* (*WISE*) has turned its infrared eyes to the sky, expecting to detect hundreds of previously unseen near-Earth objects. It will reveal the darkest members of this population, which don't reflect much visible light but do emit detectable infrared heat radiation.

Once an NEO is discovered by the ongoing surveys, previous unsuspecting observations made before the NEO discovery and follow-up observations with powerful radar telescopes can be used to refine knowledge of its future trajectory and push predictions far into the future. By bouncing radio waves off the object, and examining the return echo, the radar technique can be used to establish an exact orbit, providing us with centuries of advance notice of close encounters or impacts by NEOs.

Near-Earth Objects are catalogued at the International Astronomical Union's Minor Planet Center, located on the web at http://www.cfa.harvard.edu/iau/mpc.html. They list more than one thousand of the potentially hazardous objects, with the greatest likelihood of close approaches to Earth. The Near Earth Object website of the Jet Propulsion Laboratory, located at http://neo.jpl.nasa.gov/, carries several tables that can be sorted by close approach, date, object name, past and future approaches, and comets or asteroids. The European Spaceguard Foundation includes a Spaceguard System of observatories that are engaged in Near-Earth Object observations in order to discover them and protect the Earth from the possible threat of their collision.

the asteroid's course. If the impact is many years away and the threatening object relatively far away from us, all we have to do is give it a little nudge. By the time the asteroid reaches the Earth's vicinity, that small change in trajectory will make a big difference, enabling it to bypass the planet.

An orbiting mirror could be used to focus sunlight and vaporize the asteroid's surface, producing a jet of gas and dust that might change its course. Astronauts could rendezvous with it, attaching small explosives, rocket engines, solar sails or mass drivers that could push it into a harmless trajectory. A large solar sail would use the pressure of sunlight to move the body slowly, while a mass driver would scoop up surface material and hurl it away, creating an appropriate recoil reaction just as the small explosives or rocket engines would.

If the warning time is only a matter of months or less, the sole recourse might be to send a high-powered rocket armed with a bomb powerful enough to redirect the object or blow it up. Such a possibility has sparked the interest of bomb designers and some members of the military (Focus 13.3). A conventional nuclear weapon might be used to deflect or destroy a small, solid, rocky asteroid, but a much larger explosion could be needed to divert or pulverize a loosely bound one. Moreover, it might not be a good idea to blow the threatening object up, for the Earth might then be struck by a hail of dangerous projectiles rather than a single blast. Some of the pieces could still head straight at us. If they were not small enough to burn up in the atmosphere, the fragments might cause massive destruction in several places on the Earth.

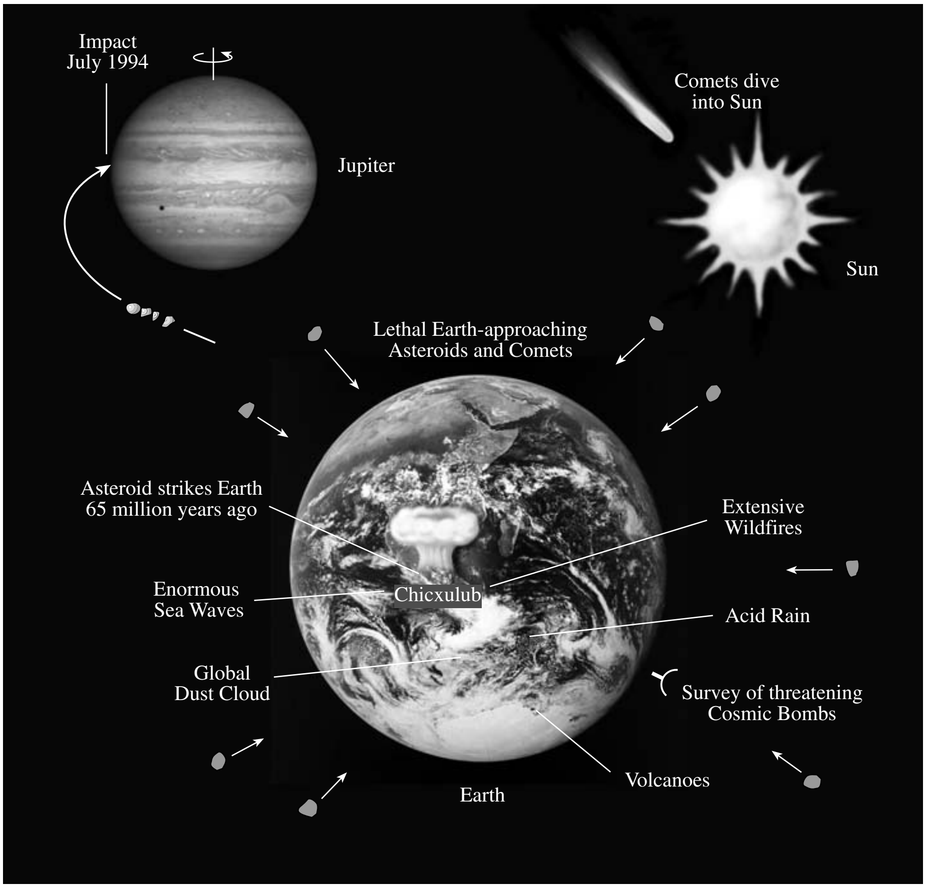


Fig. 13.16 Summary diagram

Focus 13.3 Star Wars in outer space

When it was realized that an asteroid might be hurtling toward collision with the Earth, experts in nuclear weapons and missile defense systems had something other than nuclear war to worry about. They proposed that nuclear-tipped missiles could be used to blow the threatening object up or to deflect it en masse from its earthward trajectory.

Of course, the military would want to test the warheads. With the whole world threatened you wouldn't want to use an untested weapon. In 1996, China even advocated underground tests to prepare nuclear warheads as a possible defense against the cosmic threat.

That brings up the question of who is told about plans to protect civilization from the intruder. A madman, or someone making a stupid error, could use the Star Wars weapons to deflect a previously benign asteroid into a collision course with Earth. After all, we now know how terrorists can redirect airplanes to destroy thousands of innocent lives.

In the absence of early warning and evasive action, advance knowledge of the time and place of an impact would at least allow people to store food and supplies and to evacuate regions near the impact site where damage would be the greatest. It could happen today, or it might not happen for millennia. Some of us will place our trust in God, and many of us have more immediate concerns, but it will not hurt to subscribe to a cosmic insurance policy and fully identify the swarm of cosmic objects that might be headed our way. The policy has relatively small premiums and enormous potential benefits.

14 Comets

- The sudden apparition, changing shapes, and unpredictable movements of comets have puzzled humanity for centuries. To ancient cultures they were harbingers of disaster and portents of great events.
- Comet Halley has returned to fascinate and frighten the world for more than 2000 years.
- Long-period comets, with orbital periods greater than 200 years, have been tossed into the planetary realm from a remote, spherical shell, named the Oort cloud, located about a quarter of the way to the nearest star.
- A million million (10^{12}) invisible comets have been hibernating in the deep freeze of the Oort comet cloud since the formation of the solar system 4.6 billion years ago.
- Many Jupiter family comets, with orbital periods of less than 20 years, probably came from the Kuiper belt, which lies in the outer disk of the planetary system beyond the orbit of Neptune and may contain more than a billion unseen comets.
- Comets light up and become visible for just a few weeks or months when their orbits bring them near the Sun. The solar heat then vaporizes some of the comet water ices, permitting the comets to grow large enough to be seen. The water ice sublimates, or turns directly from solid ice to water vapor.
- The solid comet nucleus is just a gigantic ball of water ice, other ices, dust and rock. Some of the comet nuclei are about the size of Paris or Manhattan and roughly one-billionth the mass of the Earth. Other comets are much smaller.
- No two comets ever look identical, and every comet changes shape and form as it whips around the Sun, but they all develop a glowing spherical cloud of gas and dust, known as the coma, when moving close enough to the Sun.
- The comet coma can be larger than the Earth and as big as the Sun, and around the coma there is an even larger envelope of atomic hydrogen, known as the hydrogen cloud, that shines in ultraviolet light.
- Some comets develop tails that flow away from the Sun, briefly attaining lengths as large as the distance between the Earth and the Sun, but other comets have no tail at all.

- Comets can have two kinds of tails: a long, straight ion tail, that re-emits sunlight with a faint blue fluorescence, and a curved dust tail that shines by reflecting yellow sunlight.
- The *Giotto*, *Deep Space 1*, *Stardust* and *Deep Impact* spacecraft have respectively peered into the icy hearts of four comets – Halley, Borrelly, Wild 2 and Tempel 1 – showing that their nuclei are blacker than coal and reflect just a few percent of the incident sunlight. *Deep Impact* continued on to encounter comet Hartley 2.
- Gas and dust jet out from the sunlit side of the nucleus of comet Halley, from fissures in its dark crust, but nearly 90 percent of the surface of its nucleus was inactive at the time of the *Giotto* encounter.
- Comet Borrelly is covered with a dark, unreflective carbon-rich material, and contains surface features that are most likely supported by solid water ice.
- Comet Wild 2 has been exposed just a few times to the Sun's intense heat, and it has a dark, pockmarked surface with pits, craters and jets of gas and dust.
- The *Stardust* spacecraft gathered dust from the coma of comet Wild 2 in January 2004, returning the dust in a capsule that was parachuted to Earth two years later. The returned comet dust contains a mix of minerals formed at both cold and high temperatures, two types of nitrogen-rich organic molecules, and the amino acid glycine.
- The *Deep Impact* spacecraft collided with comet Tempel 1 on 4 July 2005; spectroscopic examination of the ejected cloud of dust revealed fine porous material, water vapor, water ice, carbon dioxide, hydrocarbons and silicates or sand.
- The *Deep Impact* spacecraft flew past comet Hartley 2 on 4 November 2010, revealing a small active nucleus composed of two rough parts joined at a smooth waist.
- When a bright comet nears the Sun, it turns on its celestial fountains, spurting out about a million tons of water each day.
- The recoil effect of jets of matter ejected from a comet's spinning, icy nucleus can push a comet along in its orbit or oppose its motion, causing the comet to arrive closest to the Sun earlier or later than expected.
- Most of the comets seen during recorded history will vanish from sight in less than a million years, either vaporizing into nothing or leaving a black rock behind.
- About 40 000 tons of small, cosmic dust particles fall onto the Earth in a typical year, wafting gently through the atmosphere to the ground.
- Visible comets are in their death throes, but they may carry the residues of creation in their ice and dust.
- Meteor showers, commonly known as shooting stars, are produced when sand-sized or pebble-sized pieces of an icy comet burn up in the atmosphere, never reaching the ground.
- Comets strew particles along their orbital path as they loop around the Sun, and when the Earth passes through one of these meteoric streams a meteor shower occurs, recurring at the same time every year.

Fig. 14.1 The Great Comet of 1577 This drawing by a Turkish astronomer appeared in the book *Tarcuma-I Cifr al-Cami* by Mohammed b. Kamaladdin written in the 16th century. The yellow Moon, stars and comet are shown against a light blue sky. (Courtesy of Erol Pakin, Director, Istanbul Universitesi Rektorlugu.)

14.1 Unexpected appearance of comets

Every few years, on average, an unusually bright comet will blaze forth in the night sky, becoming visible to the unaided eye and sporting a graceful tail resembling long hair blowing in the wind (Figs. 14.1, 14.2). In fact, the word *comet* is derived from the Greek name *aster kometes*, meaning "long-haired stars". But a comet is not anything like a star. Their dramatic display emanates from a relatively small, blackened chunk of ice and dust, comparable to a large city in size.

Unlike the planets, the comets can appear almost anywhere in the sky, remain visible for a few weeks or months, and then vanish into the darkness. Astronomers call this period of visibility an "apparition". During its apparition, a comet changes its shape, often from night to night.

Many of the enigmatic comets travel far outside the paths of the planets and move in every possible direction around the Sun. Their orbits are inclined at all possible angles to the ecliptic, the plane of Earth's orbit, and different comets move in either the same direction around the Sun as the planets or in the opposite retrograde direction. Other comets move in tighter orbits, circling the Sun within the bounds of the outer planets.

Comets inspired awe and fear in ancient cultures. By their unexpected arrivals, these celestial intruders seemed to upset the natural order of the otherwise placid firmament, and to presage changes in the order of things on Earth, such as the death of rulers, wars and other disasters (Fig. 14.3). One comet appeared in 44 BC, the year that Julius Caesar was assassinated. The fallen emperor's adopted son declared the comet to be Caesar's soul rising to heaven, and used the apparition to gain control over the entire Roman Empire as Augustus Caesar (63 BC–14 AD). William Shakespeare (1564–1616) wrote about the comet's link to Julius Caesar's death 15 centuries later, with:

> When beggars die, there are not comets seen;
> The heavens themselves blaze forth the death of
> princes.
> Cowards die many times before their deaths;
> The valiant never taste death but once.
>
> *Julius Caesar*, Act II, Scene ii, line 30

Another famous example was the Norman conquest of England in 1066, which was coincident with the appearance of what is now called comet Halley. This comet was also seen in 1456 when the Turks conquered Constantinople, and some in Europe prayed for protection from "the Devil, the Turk and the Comet". More recently, the Great Comet of 1811 was supposed "to portend all kinds of woes and the end of the world" in Leo Tolstoy's (1828–1910) *War and Peace*.

The unexpected appearance of a bright comet was also once taken as an omen of doom and the harbinger of disaster. In *Paradise Lost*, John Milton (1608–1674) imagined a Satan-like comet that "from his horrid hair shakes pestilence and war." Even in 1910 there were speculations that comet Halley would impregnate the air with poisonous vapors such as cyanogen, and wipe out life on Earth, but there were no noticeable effects on humans or other living things when the Earth passed near the comet's tail.

Awe-inspiring comets still arrive without warning today, becoming brighter than the most brilliant stars. They are usually named after the last names of their discoverers, unless a spacecraft is involved (Focus 14.1). The brightest, most spectacular apparitions are also known as Great Comets.

At night, the Great Comets can remain visible to the unaided eye for months, and they sometimes become visible in daylight. Some of them have enormous tails (Fig. 14.4). Great Comet Hyakutake, for example, came within 0.1 AU of the Earth in 1996, with a tail that stretched one-quarter the way across the sky.

As illustrated in Table 14.1, the Great Comets all travel closer to the Sun than the Earth, which orbits at a mean distance of 1 AU, and sometimes pass quite near the Earth itself. But a decade or more can pass between the unanticipated discoveries of truly Great Comets.

Fig. 14.2 Comet Kohoutek A modern photograph of a comet's flowing tail. It was taken on 12 January 1974 with the 1.2-meter (48-inch) Schmidt telescope of the Hale Observatories with a 3-minute exposure in blue light. (Courtesy of the Hale Observatories.)

14.2 The return of comet Halley

Many of the brightest comets seem to come out of nowhere, suddenly moving past the Sun, and are never seen again. Yet, one famous comet has come back for repeat performances, fascinating the world for more than 2000 years. It is now known as comet Halley, named for the British astronomer Sir Edmond Halley (1656–1742).

Halley demystified comets by showing that at least one of them travels in an elongated orbit around the Sun. He found that the orbit of the bright comet of 1682 was similar to those of comets observed in 1607 (seen by Johannes Kepler, 1571–1630) and in 1531 (observed by Petrus Apianus, 1495–1552). All three comets moved around the Sun in retrograde orbits with a similar orientation. Halley also

Fig. 14.3 The eve of the deluge People believed for centuries that the unexpected appearance of comets was a premonition of war, death and other disasters. Here the arrival of a comet foretells the great flood at the time of Noah. The 1835 apparition of comet Halley may have influenced the artist, John Martin (1784–1854), for he finished this painting a few years later in 1840. (Collection of Her Majesty the Queen.)

Focus 14.1 Naming comets

Amateur astronomers often discover the brightest comets, diligently searching for them with small telescopes or even large binoculars. Professional astronomers sometimes accidentally come across one while using a large telescope for another purpose. In accordance with a tradition that has gone on since the time of the French comet hunter Charles Messier (1730–1813), new comets are now given the last name of their discoverer, the last names of their independent discoverers, or the acronym of a spacecraft used in the discovery, such as *SOHO*.

A prefix "P/" is now used for a periodic comet, defined to have a revolution period of less than 200 years with confirmed observations at more than one perihelion passage, and "C/" for a comet that is not periodic in this sense. A number is also added before the prefix P to designate the order of discovery. For instance, 1P/Halley was the first periodic comet known and 2P/Encke the second. Comets are also now designated by the year of their discovery, the upper-case letter identifying the half-month of the observation during that year, and a consecutive numeral to indicate the order of discovery announcement during that half-month. The letters "I" and "Z" are not used to make a total of 24 half-months. For example, the third comet reported as discovered during the second half of February 1995 would be designated 1995 D3.

knew that the Great Comet of 1456 had also traveled in the retrograde direction, and he concluded that all four apparitions were the same comet traveling along an identical elongated orbit around the Sun and appearing at 76-year intervals when coming close to our star. Halley confidently predicted its return in 1758, noting that he would not live to see it.

Halley also pioneered our understanding of cartography, diving bells, mortality tables, naval navigation, stellar proper motions, tides and trade winds, and helped with the publication of Isaac Newton's (1643–1727) powerful description of gravity and orbital motion. But Halley is best known for the comet that now bears his name, comet Halley (Fig. 14.5). It was re-discovered, on Christmas night of the predicted year.

Halley's is the most famous of the comets because it was the first to arrive on schedule. Its fame is deserved on other counts as well. It displays a complete range of comet fireworks including an exceptionally long tail, a bright head, and jets, rays, streamers and halos. Moreover, we now know that it has been observed for a longer period of time than any other comet in recorded history. The earliest apparition established with confidence from Chinese chronicles dates back to 240 BC; since then, all its perihelion passages have been retraced in the ancient or modern records of astronomers (Table 14.2).

After its 1910 apparition (Fig. 14.6), comet Halley moved away from the Sun into the outer darkness, arriving in 1948 at the remotest part of its orbit at 35 AU. The comet then turned the direction of its course, and began falling back toward the heart of the solar system with ever-increasing speed. It reached perihelion, or its closest distance from the Sun, on 9 February 1986. Comet Halley and the Earth were then on opposite sides of the Sun, so this was among the least favorable apparitions for observing the comet with the unaided

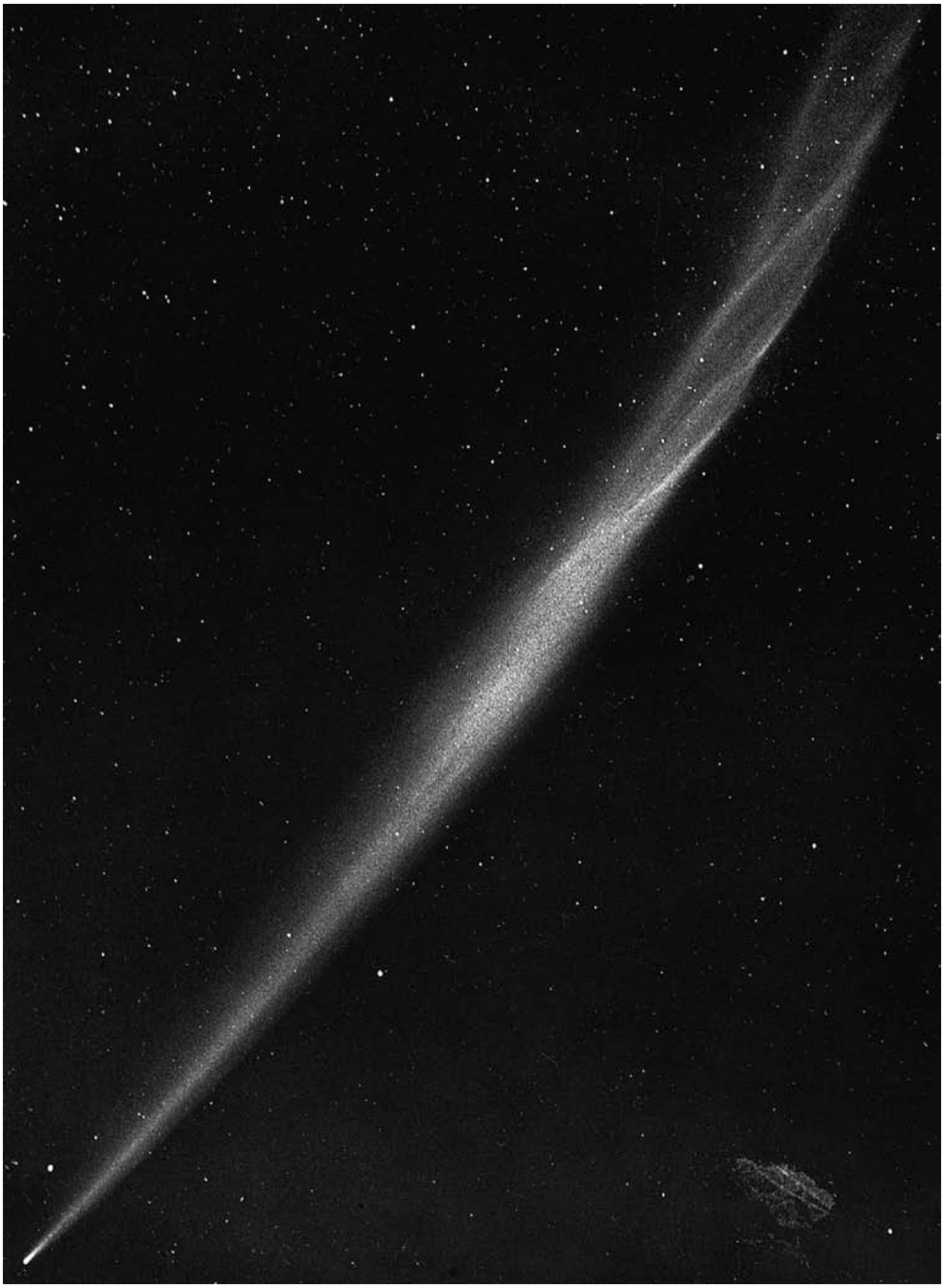


Fig. 14.4 A Great Comet lights up When a comet travels close to the Sun, the solar heat vaporizes ice from the comet's surface, and solar forces bend the liberated material into comet tails that always point away from the Sun rather than toward it. The long tail of this Great Comet stretches 120 million (1.2×10^8) kilometers, or nearly the mean distance between the Earth and the Sun, at 1 AU = 1.496×10^8 kilometers. It is also named comet Ikeya–Seki (1965 S1) after the last names of its discoverers, Kaoru Ikeya (1943–) and Tsutomu Seki (1930–), and the year and order of its discovery. (Courtesy of the Lick Observatory.)

eye. Nevertheless, it still became one of the most thoroughly studied apparitions in the history of comet research (Fig. 14.7), including visits by several spacecraft. After these visits, Halley's comet headed for the cold reaches of the solar system, to return in the Sun's neighborhood in 2061.

14.3 Where do comets come from?

Comets are primitive bodies that formed at the same time as the Sun and planets about 4.6 billion years ago. But once they come close enough to be seen, comets begin to fall apart and they must eventually vanish from sight, often in

Table 14.1 Some Great Comets of the 19th and 20th centuries[a]

Name	Perihelion date	Days visible[b]	Perhelion distance (AU)	Brightest apparent magnitude
Great Comet of 1807	19 Sept. 1807	90	0.65	1 to 2
Great Comet of 1811	12 Sept. 1811	260	1.04	0
Great March Comet of 1843	27 Feb. 1843	48	0.006	1
Comet Donati	30 Sept. 1858	80	0.58	0 to 1
Great Comet of 1861	12 June 1861	90	0.82	0 (or –2?)
Great Comet of 1865	14 Jan. 1865	36	0.03	1
Comet Coggia	09 July 1874	70	0.68	0 to 1
Great September Comet	17 Sept. 1882	135	0.008	–2
Great Comet of 1901	24 Apr. 1901	38	0.24	1
Great January Comet	17 Jan. 1910	17	0.13	1 to 2
Comet Halley (in 1910)	20 Apr. 1910	80	0.59	0 to 1
Comet Skjellerup-Maristany	18 Dec. 1927	32	0.18	1
Comet Ikeya-Seki	21 Oct. 1965	30	0.008	2
Comet Bennett	20 Mar. 1970	80	0.54	0 to 1
Comet West	25 Feb. 1976	55	0.20	0
Comet Hyakutake	01 May 1996	30	0.23	1 to 2
Comet Hale-Bopp	01 Apr. 1997	215	0.91	–0.7

[a] The perihelion distance is the distance from the Sun at the closest approach to the star, given in astronomical units (AU), roughly the mean distance between the Earth and the Sun. The apparent magnitude is a measure of the apparent brightness of a celestial object, in which brighter objects have smaller magnitudes. Sirius A, the brightest star other than the Sun, has an apparent visual magnitude of –1.5. The nearest star other than the Sun is about 0 on the magnitude scale, while Venus has an apparent magnitude of –4 when brightest, and at its brightest Jupiter appears at magnitude –2.7. Adapted from Donald K. Yeomans' *Great Comets in History*, at the website http://ssd.jpl.nasa.gov/great_comets.html.

[b] Days visible to the naked eye unaided by binoculars or a telescope.

less than a million years after first sighting. So comets are very old, but once they swing near the Sun they do not last very long.

Something must be furnishing the inner solar system with new comets, and they come from two reservoirs. One is very far away, at the fringe of the outer solar system, and another nearer one is at the edge of the planetary realm. We distinguish between these source regions on the basis of the orbital periods of the comets. Both types of small icy worlds have been hibernating in the cold outer reaches of space ever since the formation of the solar system.

The long-period and short-period comets

Traditionally, comet orbits have been classified as short period or long period with the dividing line at orbital periods of 200 years. Most newly discovered comets are long-period comets, with orbital periods larger than 200 years. They have arrived near the Sun from distant regions far beyond the major planets, traveling along very elongated trajectories that take them back to the distant regions they came from. The long-period comets are observed in the inner part of the solar system just once, arriving unannounced and unpredicted. As you might expect, they come from very far away, at the outer fringes of the solar system.

Nowadays, astronomers divide the former short-period designation into the Jupiter family, with orbital periods less than 20 years, and the Halley-type comets that have orbital periods between 20 and 200 years. The Jupiter-family comets move about the Sun with a period comparable to Jupiter's 10-year period, and have orbits affected by the gravitational influence of the giant planets.

The Jupiter-family comets are seen time and again, trapped in tight orbits within the planetary realm. It is these comets whose regular returns we are able to predict, and which we can examine in detail with spacecraft (Table 14.3). They travel on direct orbits, in the same direction as the planets orbit the Sun, and most of them have low orbital inclinations near the plane of the Earth's orbit, with mean distances from the Sun of just a few times that of the Earth.

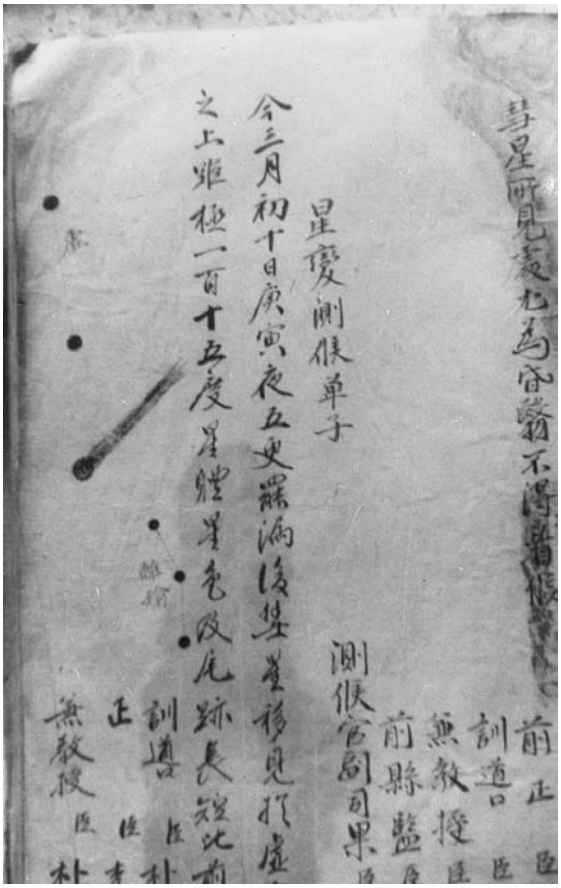


Fig. 14.5 Comet Halley in 1759 AD This Korean record of comet Halley was made during the comet's first predicted return in 1759. The Korean astronomers have been recording the appearance of comets and other unusual celestial objects for more than 3000 years. (Courtesy of Il-Seong Na, Yonsei University, Seoul.)

Table 14.2 Thirty-two perihelion passages of comet Halley[a]

240 BC	25 May
164	13 November
87	6 August
12 BC	11 October
66 AD	26 January
141	22 March
218	18 May
295	20 April
374	16 February
451	28 June
530	27 September
607	15 March
684	3 October
760	21 May
837	28 February
912	19 July
989	6 September
1066	21 March
1145	19 April
1222	29 September
1301	26 October
1378	11 November
1456	10 June
1531	26 August
1607	28 October
1682	15 September
1759	13 March
1835	16 November
1910	20 April
1986	9 February
2061[a]	28 July
2134 AD[a]	27 March

[a] The perihelion of a comet is the point in its orbit that is closest to the Sun. The future two perihelion passages of Halley's comet are predicted dates; all of the others have been recorded.

The Oort cloud of comets

The size and orientation of the trajectories of long-period comets can be explained if they come from a remote, spherical shell belonging to the outer parts of the solar system (Fig. 14.8). This vast comet reservoir is known as the Oort cloud, named after the Dutch astronomer Jan H. Oort (1900–1992) who first postulated its existence in 1950. His careful examination of the trajectories of observed long-period comets, which became visible when they entered the inner parts of the solar system, could only be explained if these comets came from a distant reservoir, which he located at between 50 000 and 150 000 AU. At greater distances the stars in the neighborhood of our solar system compete for gravitational control of the comets.

And because long-period comets enter the planetary realm at all possible angles, with every inclination to the Earth's orbital plane, they must come from a spherical shell. This would also explain the fact that long-period comets move in all directions. Roughly half of them move along their trajectories in the retrograde direction, opposite to the orbital motion of the planets.

Modern calculations suggest that there is an inner Oort cloud with an inner edge at around 3000 AU and a density falling off with greater distance. The outer Oort cloud is continuous with this, but is defined to be those objects at distances greater than 20 000 AU. The cloud fades away with increasing distance, and its tenuous outer edge is dynamically limited to about 200 000 AU by the galactic gravity field.

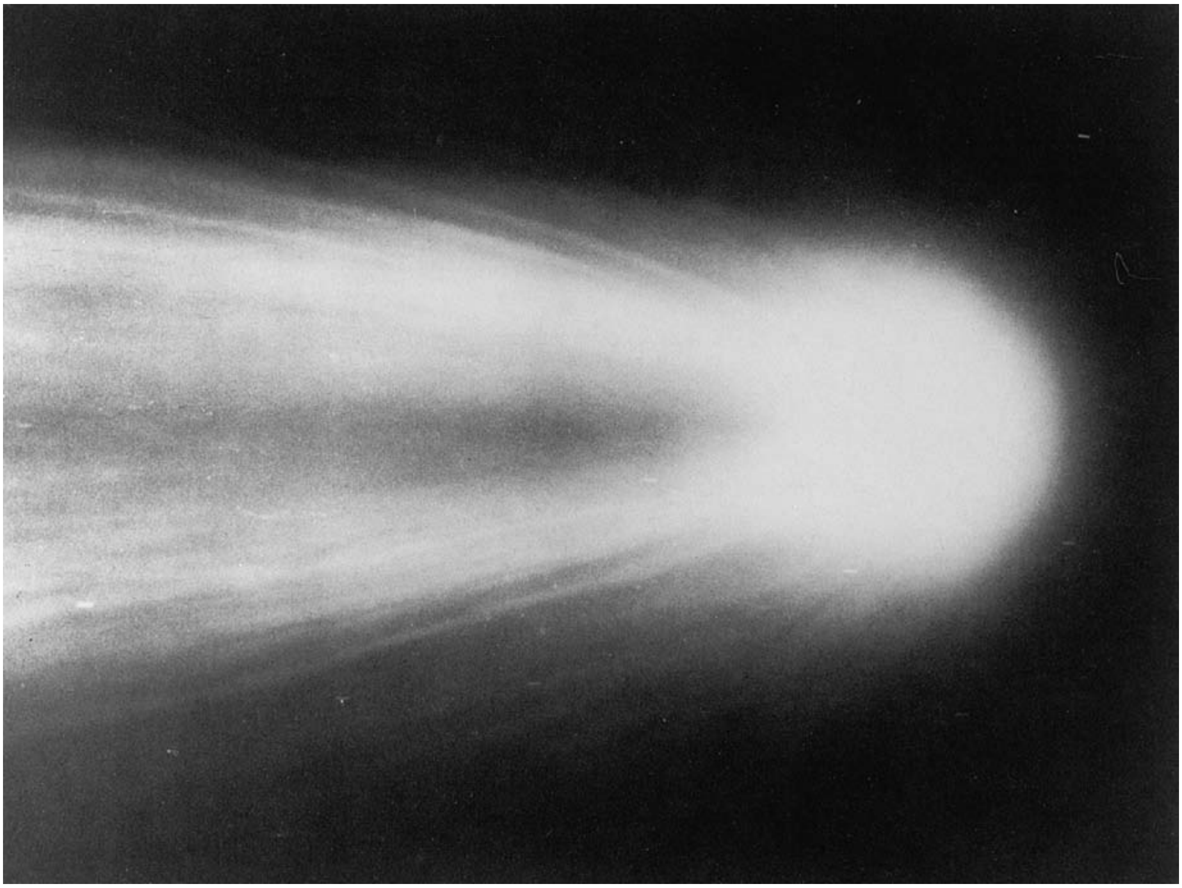


Fig. 14.6 Apparition of comet Halley in 1910 The head region or coma of comet Halley observed on 8 May 1910 with the 1.5-meter (60-inch) telescope on Mount Wilson. The comet's tail flows to the left, away from the Sun. (Courtesy of the Hale Observatories.)

The Oort cloud has also been divided into two components: the spherical outer cloud discussed by Oort and a more flattened inner cloud. The inner cloud is probably the source of the Halley-type comets; they require a closer origin to be captured into stable orbits around the Sun with periods between 20 and 200 years.

Formation of the Oort comet cloud

Where did the Oort-cloud comets originally come from? They could not have formed in their current position, because the material at such large distances from the young Sun would have been too sparse to coalesce. They instead originally condensed and agglomerated into comet-sized bodies between the orbits of Jupiter and Neptune, as the leftover bits and pieces from the formation of the solar system. Once formed, the kilometer-size comet bodies were swept out of the region by the newborn giant planets. Gravitational perturbations by the newly formed giant planets sent some of the comets to large distances from the Sun and some into the inner solar system where they faded long ago. Thus, the nascent giants acted like cosmic street-cleaners, either hurling the newborn comets into distant regions or consuming them.

Many of the comets sent to large distances escaped from the solar system. Some remained barely bound to the Sun by its weakened gravity at large distances; in a roughly spherical cloud about 10 000 to 200 000 AU in radius, the Oort cloud. Out there it takes roughly 10 million years to complete one orbit around the Sun.

Jupiter and Saturn, the two most massive planets, might have ejected some of them into interstellar space, but they also placed a significant fraction of comets into the Oort cloud. Jupiter may have additionally tossed nearby comets into a collision course with our planet, perhaps supplying some of early Earth's water and organic compounds. Uranus and Neptune, with lower masses, could not easily throw the primitive comets into the space between the stars, but they should have tossed about the same number of comets into the Oort cloud as Jupiter and Saturn did. The outer giants would have also pulled nearby comets into themselves, helping them to grow.

The comets that belong to the Oort cloud are therefore mementos of creation, frozen into the deep freeze of outer

Table 14.3 Selected short-period comets[a]

Name	Orbit period (years)	Perihelion date[b] (year)	Perihelion distance[b] (AU)	Orbital inclination (degrees)	Absolute magnitude
2P Encke	3.30	2003	0.34	11.8	9.8
6P d'Arrest	6.51	2002	1.35	19.5	8.5
9P Tempel 1[c]	5.51	2000	1.50	10.5	12.0
19P Borrelly[c]	6.88	2001	1.36	30.3	11.9
21P Giacobini–Zinner	6.61	2005	1.00	31.9	9.0
26P Grigg–Sjkellerup	5.11	2002	0.99	21.1	12.5
46P Wirtanen[c]	5.46	2002	1.06	11.7	9.0
67P Churyumov–Gerasimenko	6.57	2009	1.29	7.1	15.4
73P Schwassmann–Wachmann 3	5.34	2001	0.94	11.4	11.7
81P Wild 2[c]	6.39	2003	1.58	3.2	6.5
103P Hartley 2	6.46	2004	1.05	13.6	16.6

[a] Adapted from Kenneth R. Lang, *Astrophysical Data: Planets and Stars* (New York, Springer Verlag, 1992) and Gary M. Kronk's comet website http://cometography.com.
[b] The given perihelion date is the first to occur in the 21st century, and the perihelion distance is in AU, the mean distance between the Earth and the Sun. The absolute magnitude is a measure of the intrinsic brightness of a comet, and a smaller magnitude indicater a brighter comet.
[c] Comets that have either been visited by spacecraft in the past or have been considered for encounters in the future, listed in the order of their recognition, or periodic comet number.
[d] Also asteroid (4015).

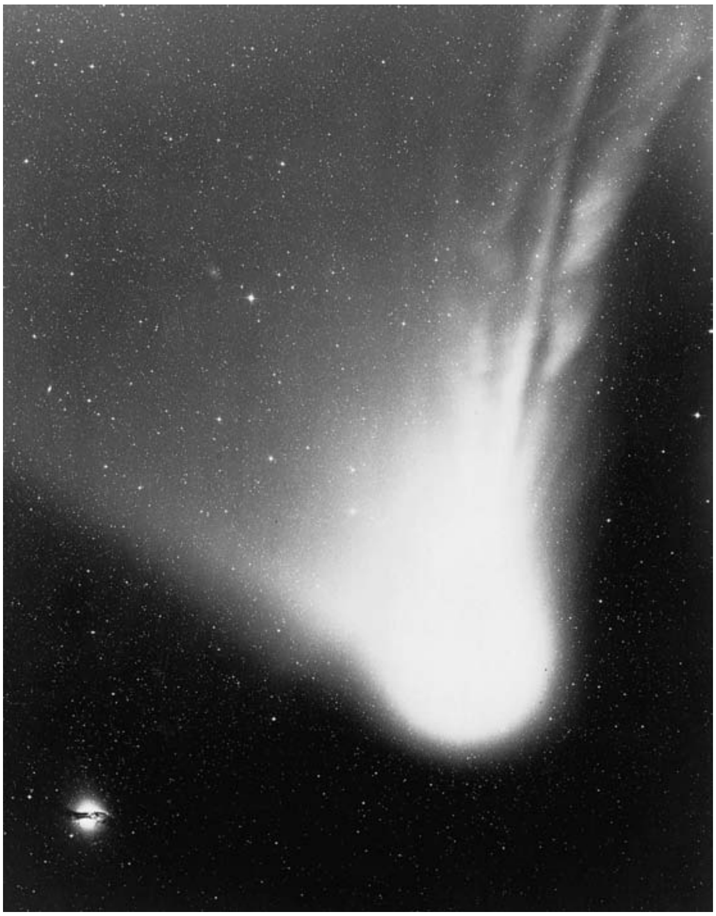


Fig. 14.7 The return of comet Halley in 1986 Rays, streamers and kinks can be seen in the ion tail of comet Halley during its 1986 return to the inner solar system. The broad, fan-shaped dust tail can also be seen. The radio galaxy known as Centaurus A, or NGC 5128, can be seen in the bottom left corner. It is about ten trillion, or 10^{13}, times further away from the Earth than the comet is. Photograph taken by Arturo Gomez on 15 April 1986 with the Curtis Schmidt telescope of Cerro Tololo. (Courtesy of NOAO.)

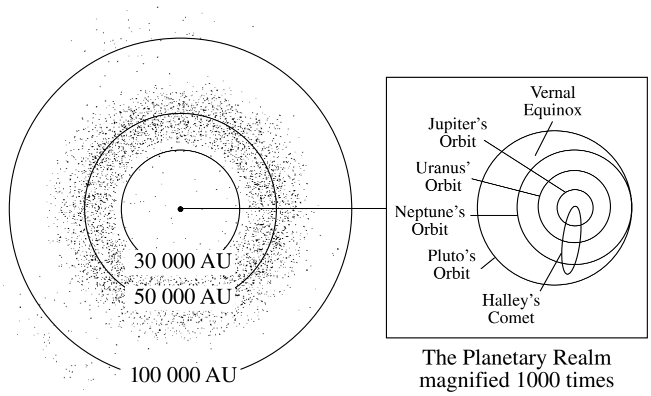


Fig. 14.8 The Oort comet cloud More than 200 billion comets hibernate in the remote Oort comet cloud, shown here in cross-section. It is located in the outer fringes of the solar system, at distances of about 100 000 AU from the Sun. By comparison, the distance to the nearest star, Proxima Centauri, is 270 000 AU, while Neptune orbits the Sun at a mere 30 AU. The planetary realm therefore appears as an insignificant dot when compared to the comet cloud, and has to be magnified by a factor of 1000 in order to be seen. This comet reservoir is named after the Dutch astronomer Jan H. Oort (1900–1992) who first postulated its existence in 1950.

space, tumbling unseen in the remote blackness for billions of years. They are much too small and too far away to be seen. So they will remain forever invisible, and will never be directly detected. As the Stoic philosopher Seneca (4 BC to 65 AD) put it:

> How many bodies besides these comets move in secret, never rising before the eyes of men? For God has not made all things for man.
>
> *Natural Questions, Book 7, Comets*

A few of them occasionally return as comets that we see.

Dislodging comets from the Oort cloud

But how do comets fall from the Oort comet cloud to the heart of the solar system? At such enormous distances, comets in the Oort cloud are loosely bound by the Sun's gravity, and are easily disturbed by massive objects passing by, such as nearby stars or interstellar molecular clouds, which throw some of the comets back into the planetary system. The random gravitational jostling of individual stars passing nearby, for example, knocks some of the comets in the Oort cloud from their stable orbits, either ejecting them into interstellar space or gradually deflecting their paths toward the Sun. Every one million years, about a dozen stars pass close enough to stir up the comets, sending a steady trickle of comets into the inner solar system on very long elliptical orbits. A giant interstellar molecular cloud can also impart a gravitational tug when it moves past the comet cloud, helping to jostle some of them out of their remote resting-place. Tidal forces generated in the cloud by the disk of our Galaxy, the Milky Way, also help to feed new long-period comets into the planetary region.

As time goes on, the accumulated effects of these tugs will send a few comets in toward the Sun – or outward to interstellar space. If the several hundred new comets observed during recorded history have been shuffled into view by the perturbing action of nearby stars or molecular clouds, then there are at least 100 billion, or 10^{11}, comets in the Oort cloud. There may be a trillion, 10^{12}, or even 10 trillion, 10^{13}, of them. This large population of unseen comets can sustain the visible long-period comets and persist without serious depletion for many billions of years, until long after the Sun expands to consume Mercury and boil the Earth's oceans away.

The Kuiper belt of comets

The Jupiter family of comets, with orbital periods less than 20 years, cannot come from the Oort cloud. These comets have relatively small orbits tilted only slightly from the orbital plane of the Earth, and they usually move in the same prograde direction as the planets. Unlike their longer-period cousins, the motions of the Jupiter-family comets resemble those of the planets, and their origin requires a close-in, flattened source.

This is now thought to be the Kuiper belt (Fig. 14.9), a ring of small icy objects at the outer edge of the planetary realm, just beyond the orbit of Neptune and a thousand times closer than the Oort cloud. It is named after the Dutch-American astronomer Gerard P. Kuiper (1905–1973) who predicted its existence in 1951. The name Edgeworth–Kuiper belt is used in the United Kingdom, acknowledging Kenneth E. Edgeworth's (1880–1972) proposal of the belt's existence in 1943.

Kuiper believed that comets must have formed throughout the early solar system. Although the giant planets cleared out any comets that were in their vicinity, the comets that formed beyond these planets should still be there. The density in this outer region of the primeval planetary disk was so low that the small objects did not coalesce into a single larger planet. They instead formed the flattened Kuiper belt of 100 million to 10 billion, or 10^8 to 10^{10}, small frozen worlds that have remained there for billions of years.

Armed with sensitive electronic detectors and powerful telescopes, astronomers have shown that the planetary system does not end abruptly at Neptune. They have discovered a substantial population of small, previously unseen

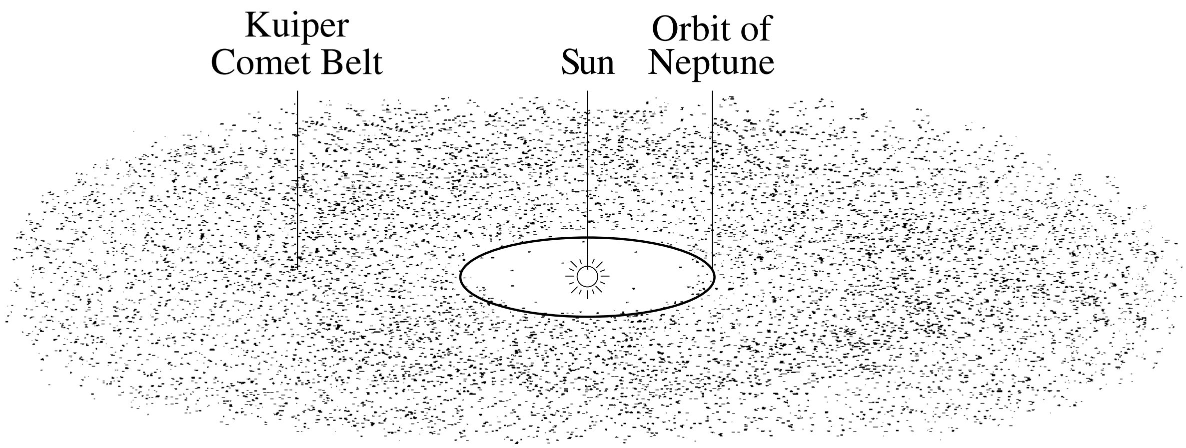


Fig. 14.9 The Kuiper belt of comets A repository of frozen, comet-sized worlds resides in the outer precincts of the planetary system, just beyond the orbit of Neptune and near the orbital plane of the planets. Known as the Kuiper belt, it is thought to contain 100 million to 10 billion comets. Many Jupiter-family comets are tossed into the inner solar system from the Kuiper belt. It is named after the Dutch-American astronomer Gerard Kuiper (1905–1973).

bodies in the Kuiper belt, each millions of times fainter than can be seen with the unaided eye (Chapter 15). All of these newly found dwarf planets travel in trans-Neptunian orbits that are only slightly tilted from the ecliptic, encircling the planetary system somewhat like the ring that wraps around Saturn. In fact, Pluto is probably one of them. Based on these detections, scientists estimate that the Kuiper belt contains tens of thousands of objects larger than 100 kilometers in size.

All of the objects observed in the Kuiper belt are larger than comets, which have a nucleus of about 10 kilometers across, but this is an observational selection effect. Smaller objects cannot be directly seen in the Kuiper belt with existing telescopes and detectors. The belt ought to contain a large population of small comet-sized bodies, just a few kilometers across. There are probably at least a billion pristine comets located just beyond the orbit of Neptune, each so faint and distant that they cannot be seen.

Unlike the comets in the Oort cloud, those in the Kuiper reservoir formed at their current locations at the dim horizon of the planetary realm, and have not been significantly perturbed since the origin of the solar system about 4.6 billion years ago. But Neptune's gravity slowly erodes the inner edge of the Kuiper belt, within about 45 AU from the Sun, launching comets from that zone into the inner solar system.

A few of the small Kuiper-belt objects are routed from their reservoir by the gravitational influence of Neptune and the other giant planets, and become regularly appearing Jupiter-family comets of shorter orbital periods. Like the comets in the Oort cloud, however, most of the comets in the Kuiper belt never come anywhere near the Sun.

14.4 Anatomy of a comet

All of the comets in the Oort cloud, and most of those in the Kuiper belt, are invisible. They are the nuclei of comets that can only be seen if they are dislodged and sent near the Sun. Each nucleus is the solid, enduring part of a comet. It is just a gigantic ball of frozen water ice and other ices laced with darker dust and pieces of rock, only a few kilometers in size. Light from the distant Sun is much too feeble to warm the comet ices, which remain frozen solid at the low temperatures in the remote comet reservoirs, so the comet never changes in size out there.

Unlike the planets, a comet lights up and becomes visible for a brief, fleeting interval during its long journey through space. It can often be detected only when it moves into the inner solar system, within the orbits of the terrestrial planets (Focus 14.2). Then the comet loops around the Sun and heads outward in more or less the same direction that it came from. As it moves away from the Sun, a comet receives less solar heat, becoming cold and inert and fading into darkness.

No two comets ever look identical (Fig. 14.10), just as no two snowflakes are alike, but most comets have basic features in common. When they emerge from the deep freeze of outer space and move toward the Sun, the comets then become visible as an enormous moving patch of light. This glowing, misty ball of light is called the *coma*, the Latin word for "hair". One or more tails can eventually stream from the coma, in a direction away from the Sun.

Comet gas and dust are initially ejected primarily in the general direction of the Sun; solar forces push them into tails that flow in the opposite direction. As a result, a

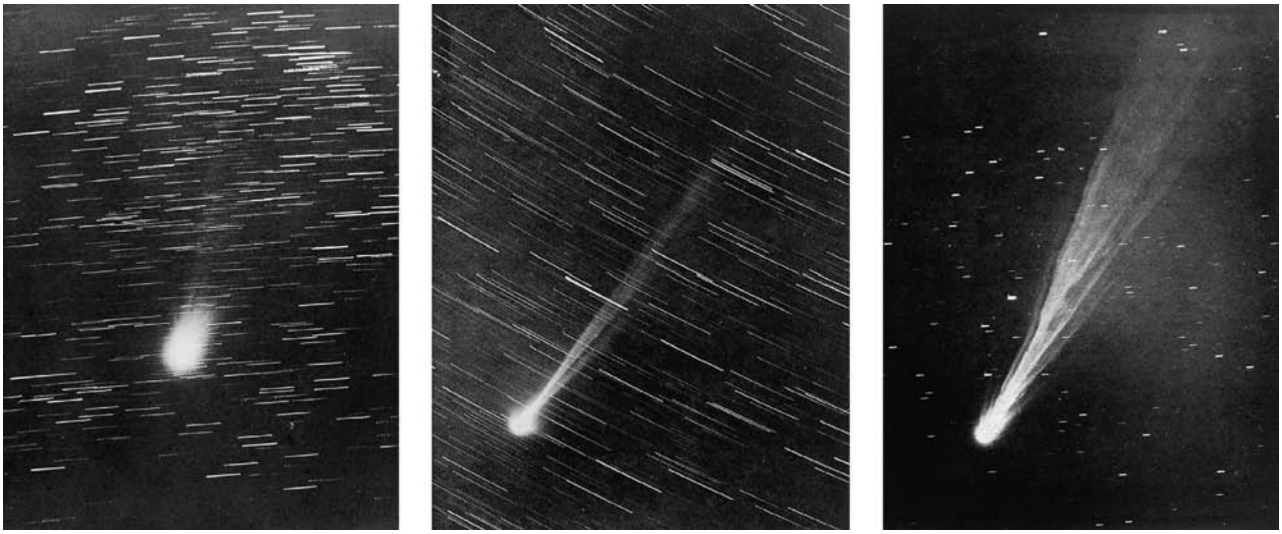


Fig. 14.10 Three kinds of comet shapes Comet Perrine (1902 III) shows a transparent coma and tail (*left*), comet Finsler (1937 V) exhibits a coma and tail that are asymmetrical (*center*), and comet Morehouse (1908 III) is remarkable for the rapid variations in the structure of its tail (*right*). When a telescope follows a comet, stars move across the field of view, producing numerous short star trails. [Courtesy of the Royal Observatory Greenwich (*left and right*) and the Norman Lockyer Observatory (*center*).]

Focus 14.2 What turns a comet on?

When a comet nucleus emerges from the deep freeze of outer space and moves toward the Sun, the increased solar heat causes the comet's surface material to sublimate, with gases escaping through fissures in the crust of the nucleus. At the low-pressure conditions of space, the solid ice goes directly into gas without passing through a liquid state, in a process called sublimation, just as dry ice does on Earth. The escaping gases also carry along dust particles. The gas and dust make the comet grow in size, enabling it to be seen.

The distance at which an invisible comet nucleus turns on, and grows large enough to be seen, varies from comet to comet. Some Jupiter-family comets are first seen at several astronomical units from the Sun, but new long-period comets that are traversing the planetary system for the first time can be detected at greater distances. For instance, many first-time visitors to the solar neighborhood become unusually bright and extensive at distances of 5 AU or more.

The outer layers of old comets, which have made many passages close to the Sun, have been "cooked" and partially stripped off, leaving behind a dark insulating crust composed largely of dust. Solar radiation has a more difficult time penetrating this material than the fresh, icy surface of a new comet. This explains the limited loss of material from periodic comets that have been repeatedly exposed to the Sun.

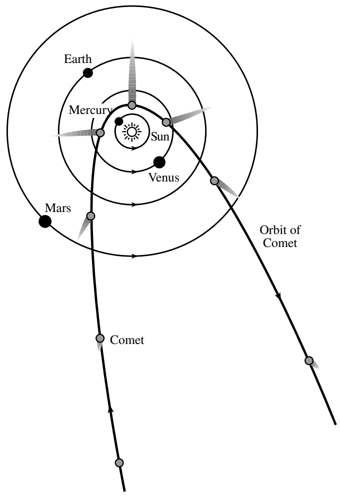


Fig. 14.11 Trajectory and tails of a comet The path and changing shape of a typical comet as it enters the inner solar system. Note that the tail of the comet is oriented away from the Sun, independent of the direction of travel of the comet.

comet travels headfirst when approaching the Sun and tail first when moving away from it (Fig. 14.11).

A comet exerts very weak gravity, so the gas and dust that are blown off the surface of the central nucleus easily escape to interplanetary space, forming the coma and comet tails. All this material is lost to the comet forever, and must be continuously replenished from the solid comet nucleus.

The visible coma, or head, is a spherical cloud of gas and dust that has emerged from the nucleus, which it surrounds like an extended atmosphere. The central body, the

Table 14.4 Structural features of a comet[a]

Feature	Size	Composition	Appearance
Nucleus	1 to 10 kilometers	Dust, ice and rock	Very dark
Coma	Up to 0.01 AU	Neutral (un-ionized) molecules and dust	Slightly yellow
Hydrogen cloud	Up to 0.1 AU	Hydrogen atoms	Ultraviolet radiation
Dust tail	Up to 0.1 AU	Dust particles	Yellow, curved
Ion tail	Up to 1 AU	Ionized molecules	Blue, straight

[a] One AU, roughly the average distance between the Earth and the Sun, is about 149.6 billion, or 1.496×10^8, kilometers.

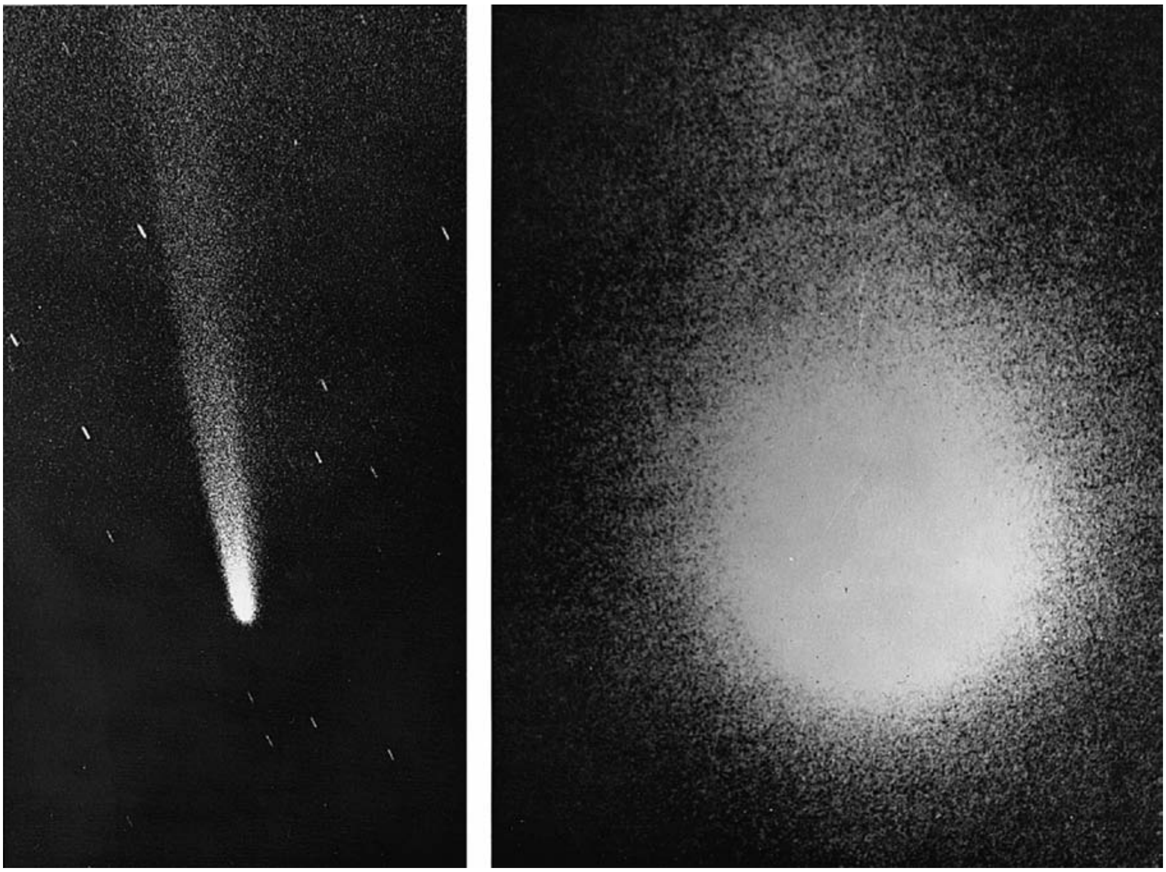


Fig. 14.12 Hydrogen cloud of a comet A comparison of the visible image (*left*) of comet Kohoutek with a far ultraviolet image (*right*) on the same scale, taken from *Aerobee* rocket flights on 4 and 7 January 1974. The ultraviolet image shows a gigantic cloud of hydrogen nearly 10 million kilometers in size, or about 10 times bigger than the Sun. It is being fed by the comet nucleus at the rate of 500 billion billion billion, or 5×10^{29}, atoms of hydrogen every second. The large size of the hydrogen cloud is due to the fact that hydrogen atoms are much lighter, and move into space faster, than the other atoms, ions, molecules and dust particles, which produce the visible light of the coma. (Courtesy of Chet B. Opal, NRL.)

nucleus, has typical dimensions of 1 to 10 kilometers, and the bulk of its composition is water ice and dust. The coma sometimes reaches a million kilometers in size, which is about as large as the Sun, and they usually become bigger than the Earth (Table 14.4).

A vast cloud containing hydrogen atoms that emit ultraviolet radiation envelops the coma and nucleus. Observations of this glow – invisible to the eye – indicate that the hydrogen halo can be 10 million kilometers across (Fig. 14.12), or about 10 times bigger than the Sun. The atomic hydrogen is produced when water molecules, released from the comet nucleus, are torn apart by energetic sunlight. The relatively light hydrogen atoms travel at high speeds to great distances before they are also

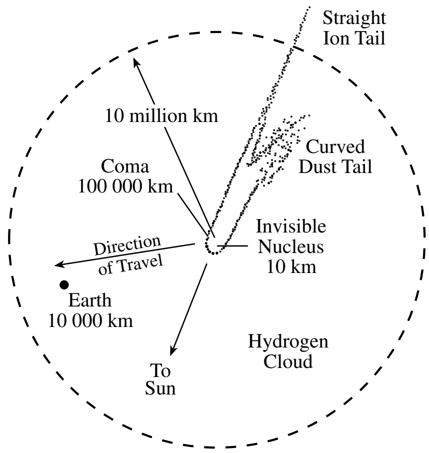


Fig. 14.13 Anatomy of a comet What you see when looking at a comet depends on how you look at it. The nucleus of a comet is usually invisible, unless a spacecraft is sent in to take a glimpse. A comet first becomes visible when it develops a coma of gas and dust. When the comet passes closer to the Sun, long ion and dust tails become visible, streaming out of the coma in the direction opposite to the Sun. When looking at a comet in ultraviolet light, the hydrogen atoms in its huge hydrogen cloud are detected.

ionized by the Sun's energetic light and swept away by its winds.

If the comet travels closer to the Sun than about 1.5 AU, it usually develops a gossamer tail; but many comets that stay outside this distance have no tail. The wispy, ghost-like tails are paler and more tenuous than the coma. The long flowing tails sweep across the sky in regal splendor, attaining lengths of 10 million, or 10^7, kilometers, and even 100 million, or 10^8, kilometers, about 1 AU. Thus, the tails of comets can briefly become the largest structures in the solar system.

Yet the comet tails appear much more substantial than they really are. You can sometimes see stars shining through the tails as if they were not there at all. It is therefore no wonder that the Earth has passed through many comet tails unscathed. So comet tails look awesome, but they contain so little matter that they are very close to being nothing at all.

To sum up, a comet's anatomy consists of a concealed nucleus, an Earth-sized or Sun-sized coma, a vast hydrogen cloud, and two types of tails, the dust and ion tails (Fig. 14.13). But a comet's anatomy is not a static thing, for comets are always changing shape. All of the comet tails grow when the comet approaches the Sun, and shrink when the comet moves away from the Sun. There is no such thing as a typical comet tail. They differ in shape, size and structure. Some comets have multiple tails, some have only one tail, and others have no tail.

14.5 Two comet tails

Some comets show two types of tails at the same time. They are the long, straight, blue ion tails, attributed to gas escaping from the comet, and the shorter, curved yellow dust tails, composed of small solid particles driven away from the comet as its ices sublimate and the gases expand. Most of the dust is comparable in size to the width of a human hair, but larger solid particles are also released from the comet, with sizes comparable to sand or even pebbles.

The gases liberated by a comet nucleus become ionized by the action of solar ultraviolet radiation and emit the faint blue light of the ion tail by fluorescence. The dust tail shines only by reflecting yellow sunlight. Since the individual dust particles enter slightly different orbits of their own, the dust tail often spreads out into a fan shape (Fig. 14.14). A comet may have a dust tail, an ion tail, both types of tail, or no tail at all.

But what are the solar forces that blow the gas and dust into comet tails? The gentle pressure of sunlight pushes the tiny, solid dust grains along curved paths as the comet moves through space. When the Sun's light bounces off the dust particles, it gives them a little outward push, called radiation pressure, and this forces them into the dust tails. For larger solid particles, comparable in size to sand or pebbles, the Sun's gravitational pull overcomes the radiation pressure, and so these particles stay near the orbital path of the comet and they do not enter the dust tails.

A solar wind of electrically charged particles and magnetic fields propels and constrains the ions on straight paths away from the Sun. The solar wind, which continuously flows away from the Sun, also accelerates the ions to high velocities.

Thus, the ion tail acts like a windsock and, in fact, the existence of the solar wind was hypothesized from observations of comet ion tails before the age of space exploration. The gas lost from a comet is ionized by ultraviolet sunlight, producing an ionosphere that envelops the comet nucleus. Magnetic fields carried by the solar wind are unable to penetrate the ionosphere, so they pile up in front of it and drape around it to form nearly parallel, adjacent magnetic field lines that point toward and away from the Sun. Guided and constrained by these folded magnetic field lines, the comet ions are pushed away from the Sun by the much faster solar wind particles, forming a straight, blue ion tail.

But the interplanetary magnetism extending from the Sun is divided into sectors that point in opposite directions, toward and away from the star. When a comet crosses from one sector to another, the magnetism that envelops its ion tail becomes pinched and the comet loses the tail,

Fig. 14.14 Dust tail of a comet This photograph of comet West (1976 VI) shows a broad, curved, pearly-hued dust tail. Because dust particles scatter sunlight, the dust tail has a slightly yellow color. It has a delicate lacy structure, created by countless dust particles shed from the comet nucleus over many days. (Courtesy of Stephen Larson, LPL/U. Arizona.)

somewhat like a tadpole. Unlike a tadpole, the comet soon grows another ion tail.

14.6 Spacecraft glimpse the comet nucleus

There is an invisible source of everything we can see when a comet passes near the Sun. It is the central nucleus of a comet, a small, icy object, typically 1 to 10 kilometers across, which is hidden within the brilliant glare of the coma's fluorescing gases and reflected sunlight.

A comet nucleus cannot be resolved using conventional telescopes, even when the comet passes very near to the Earth. A comet nucleus at a distance of only 0.2 AU would subtend an angle of about 0.002 seconds of arc. Ground-based telescopes are limited by atmospheric turbulence to a resolution of about 1.0 seconds of arc, and even outside our atmosphere the *Hubble Space Telescope* has a resolution of about 0.1 seconds of arc. So the only way to see the normally invisible comet is to send a spacecraft in close, to see through the coma and down to the detailed surface of the nucleus.

Table 14.5 Imaging missions to comets

Spacecraft	Comet	Encounter date
VEGA 1	Halley	6 March 1986
VEGA 2	Halley	9 March 1986
Giotto	Halley	14 March 1986
Deep Space 1	Borrelly	22 September 2001
Stardust	Wild 2	2 January 2004
	Tempel 1	14 February 2011
Deep Impact	Tempel 1	4 July 2005
	Hartley 2	4 November 2010
Rosetta[a]	Churyumov–Gerasimenko	August 2014

[a] The *Rosetta* spacecraft is on its way to a future encounter with a comet in 2014.

Imaging missions to comets are summarized in Table 14.5. They began by encountering comet Halley in March 1986, and continued with three Jupiter-family ones, comet Borrelly in September 2001, comet Wild 2 in

January 2004 and comet Tempel 1 in July 2005. The imaging is expected to continue in 2014 when the *Rosetta* spacecraft encounters comet Churyumov-Gerasimenko.

As we shall see, all of the comet space missions confirm the view that the nucleus of a comet is a single, solid frozen mixture of ices laced with rocky dust. The nucleus is made up mostly of water ice, but other ices such as carbon dioxide and methane are also present. As it approaches close to the Sun, energy supplied by solar radiation raises the temperature of the near-surface layers facing the Sun, and sublimation of ices, mostly water ice, takes place. The water ice goes directly from solid to the gaseous state, or to water vapor, without passing through the liquid state. This sublimation of ices produces the gas molecules that form the gas coma and subsequently the ion tail. When the ices sublimate, the embedded dust particles are also released to form the dust tail. The dust particles that are not carried away and fall back onto the nucleus form a dark insulating crust on the surface.

Observing the nucleus of comet Halley

An international flotilla of six spacecraft, belonging to four space agencies, flew by comet Halley in March 1986, to examine the gas and dust in the vicinity of the comet and to photograph its nucleus. Japan launched the *Sakigake* and *Suisei* spacecraft, meaning "pioneer" and "comet", which observed the comet from a safe distance and measured the interaction of the solar wind with the comet's atmosphere. The American probe *International Cometary Explorer* (*ICE*) also examined the solar wind upstream from the comet. *ICE* had already flown through the tail of the short-period comet 22P/Giacobini-Zinner on 11 September 1985. The two Soviet probes, named *VEGA 1* and *VEGA 2*, penetrated to within 9000 kilometers of the sunlit side of the nucleus, and the European *Giotto* approached to within 596 kilometers.

The major objective of the cameras on board both the two *VEGAs* and *Giotto* was to image the bare surface of a comet nucleus, which no one had ever seen. All three spacecraft penetrated the coma of comet Halley and detected its nucleus, despite damage by the hail of comet particles.

Giotto obtained the best images, with the highest resolution (Fig. 14.15), showing that the nucleus of comet Halley has an elongated, non-spherical and irregular shape with dimensions of $16 \times 8.5 \times 8.2$ kilometers, about the size of Paris or Manhattan. For a mass density about that of water, this volume corresponds to a mass of about 10^{15} kilograms or 1000 billion tons. A varied, lumpy topography was seen, with craters, valleys, hills, and mountains. Bright jets were spewing gas and dust from the comet's sunlit side, but there wasn't much white ice in site.

The ice had been evaporated away from the outer layers of the nucleus, to leave a dark tar-like crust all over the surface, reflecting only about 4 percent of the sunlight falling on it. When Halley passes close to the Sun, and the solar heat makes the comet active, gas jetting out from cracks and holes in its surface carries dust with it. But when the comet is far away from the Sun at the other end of its orbit, some of the black dust settles back down over the surface to form a dark crust, rich in carbon compounds.

Gas and dust can now only get out in vents where the crust has broken to expose the underlying ice, rather than from the whole sunlit hemisphere. The sunward jets are emitted from roughly 10 percent of the total surface area. Nearly 90 percent of the surface was inactive at the time of observation.

When the nucleus of comet Halley was at its maximum rate of gas emission, near its closest approach to the Sun, it released up to 1.7×10^{30} water molecules per second, or up to 56 tons of water every second. On average, the comet nucleus was emitting 20 tons of gas every second during the *Giotto* encounter, and about 80 percent of that was water. This is comparable to a loss of about 1.4 million tons of water every day, the approximate amount of water needed to supply the hydrogen halos observed around other comets. The nucleus of comet Halley was also ejecting about 10 tons of dust every second. This fully confirmed the icy conglomerate model in which a comet nucleus consists mainly of water ice and dust particles.

If the comet suffers such a high rate of water loss during the few months when it is close to the Sun, then it must lose on the order of 100 million tons of water ice during the course of each orbit. With a total mass of 100 billion tons, the comet can survive for about 1000 orbits, or 76 000 years, before it wastes away.

Deep Space 1 encounters comet Borrelly

Comet Borrelly also has a tar-black surface, as unreflective as that of comet Halley. The *Deep Space 1* spacecraft was directed toward this comet after completing its primary mission of flight-testing an ion engine and other advanced technologies. On 22 September 2001, the spacecraft whizzed by comet Borrelly at a distance of just 2200 kilometers, revealing an irregular chunk of rock and ice, about 8 kilometers long and perhaps 4 kilometers wide (Fig. 14.16). It is covered with a dark carbon-rich slag that reflects only about 3 percent of the incident sunlight, on average; comparable to the reflectivity of the powdered toner used in laser printers.

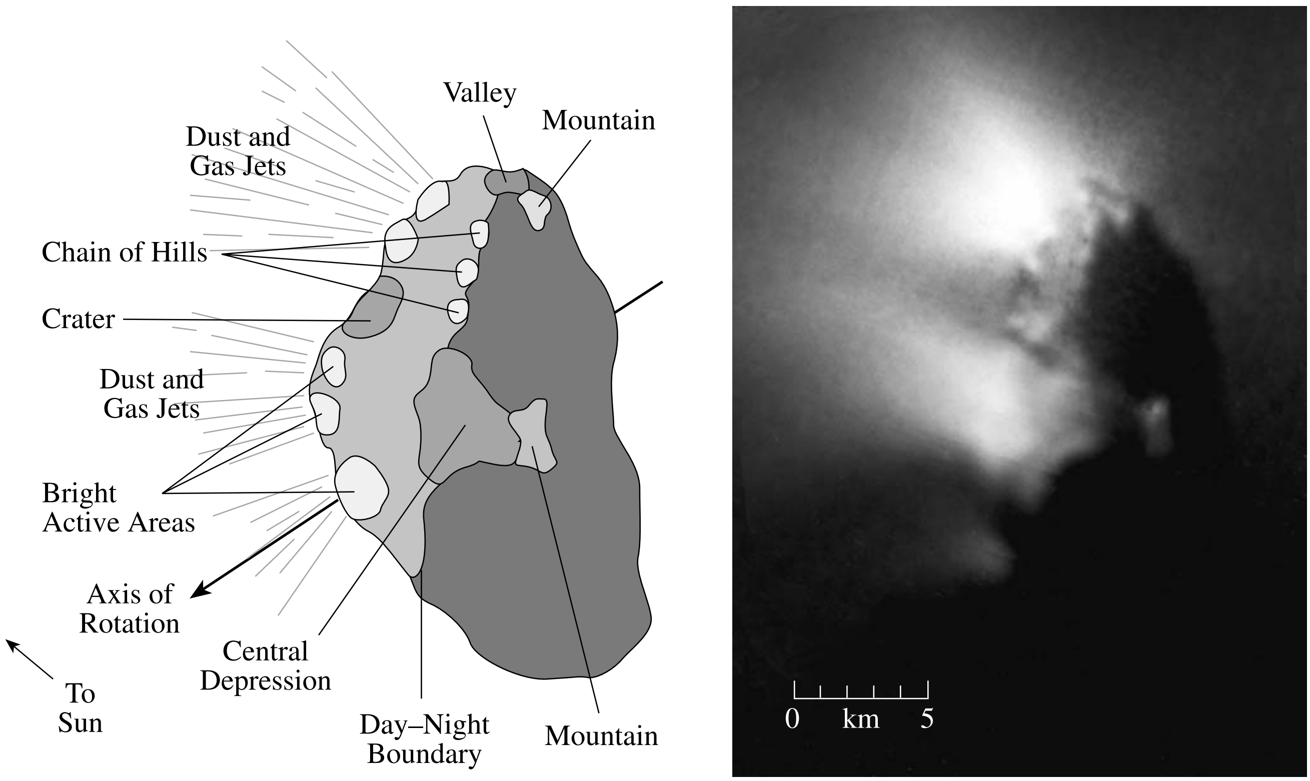


Fig. 14.15 Nucleus of comet Halley A composite image of the nucleus of comet Halley (*right*) obtained using images taken in March 1986 with the camera on board the *Giotto* spacecraft, from a distance of 6500 kilometers before comet dust destroyed the camera. It is compared with a schematic drawing (*left*) that highlights the major features recognizable in the photographs. The nucleus is about 16 kilometers long and 8 kilometers wide. Dust and gas geyser out of narrow jets from the sunlit side of the nucleus, but about 90 percent of the surface is dark and inactive. The gas is mainly water vapor sublimated from ice in the nucleus, while a significant fraction of the dust may be dark carbon-rich matter. A dark surface crust, which insulates most of the underlying ice, is blacker than coal, reflecting about 4 percent of the incident sunlight. "Mountains" rise about 500 meters above the surrounding terrain, while a broad "crater" is depressed about 100 meters. (Image courtesy of Harold Reitsema of the Ball Aerospace Corporation and Horst Uwe Keller.)

The surface of this comet nucleus has no water ice in sight, for it has all been covered up during the comet's successive passages near the Sun, leaving the dark crust behind like a dirty street a few days after a snowstorm. The sublimation has also produced a variety of surface features likened to ridges, hills, depressions, mesas and deep fractures. Beneath the rugged terrain and insulating dust there must be a supporting nucleus of solid water ice.

Stardust brings some of comet Wild 2 back home and is retargeted to comet Tempel 1

The *Stardust* spacecraft was launched on 7 February 1999, for an encounter with comet Wild 2 in early 2004. This comet is a fairly recent arrival to the inner solar system, captured into its current orbit by a close encounter with Jupiter in 1974. So it has traveled near the Sun only a few times, and its material has probably not been significantly altered by solar heat and extensive sublimation, when compared to comet Halley that has made hundreds and perhaps thousands of close passes by the Sun.

The flyby of comet Wild 2 on 2 January 2004 came within 236 kilometers of the nucleus, and resulted in detailed high-resolution images. They revealed a rugged diverse surface terrain on a dark, round body about 4 kilometers across and reflecting only about 3 percent of the incident sunlight (Fig. 14.17). Steep-walled pits and craters with flat floors were found beneath violent jets of gas and dust, shooting out from the internal, solid water ice required to feed the jets and support the surface.

The main goal of the *Stardust* mission was to collect samples of comet dust particles during the flyby and bring them home to Earth. The pioneering spacecraft caught the pristine comet dust by extending a paddle covered with an exotic substance called aerogel. It has one of the lowest

Fig. 14.16 Nucleus of comet Borrelly A camera on board the *Deep Space 1* spacecraft peered into the icy heart of comet Borrelly on 22 September 2001, taking this image from a distance of 3400 kilometers. The nucleus is shaped like a gigantic bowling pin, with a length of about 8 kilometers and a width of roughly half that size. A dark veneer of material covers most of the nucleus, reflecting only about 3 percent of the incident sunlight on average. Rugged terrain is found on both ends of the nucleus, while bright smooth plains are present in the middle. Jets of gas and dust shot out from all sides of the comet's nucleus as it rotated, producing a flow of ions that was not centered on the nucleus. (Courtesy of NASA/JPL.)

mass densities known, essentially because it is filled with holes, and was thus able to capture the dust intact. The dust particles would have evaporated when striking a denser, metallic surface.

The collected dust samples were placed within a capsule that was carried back to Earth and parachuted to the desert salt flats in Utah on 15 January 2006, while the rest of the spacecraft remained in space. The sample return capsule exceeded all expectations, containing thousands of coma particles that have been analyzed and scrutinized in terrestrial laboratories (Focus 14.3).

The *Stardust* spacecraft has been reused to encounter comet Tempel 1 on 14 February 2011, providing unique observations of a comet nucleus after its close approach to the Sun. This is a second visit to the comet, which was previously struck by the *Deep Impact* spacecraft on 4 July 2005.

Deep Impact strikes comet Tempel 1 and encounters comet Hartley 2

The *Deep Impact* spacecraft was launched on 12 January 2005 on a six-month journey to comet Tempel 1. The mission consisted of two parts, a flyby spacecraft, which

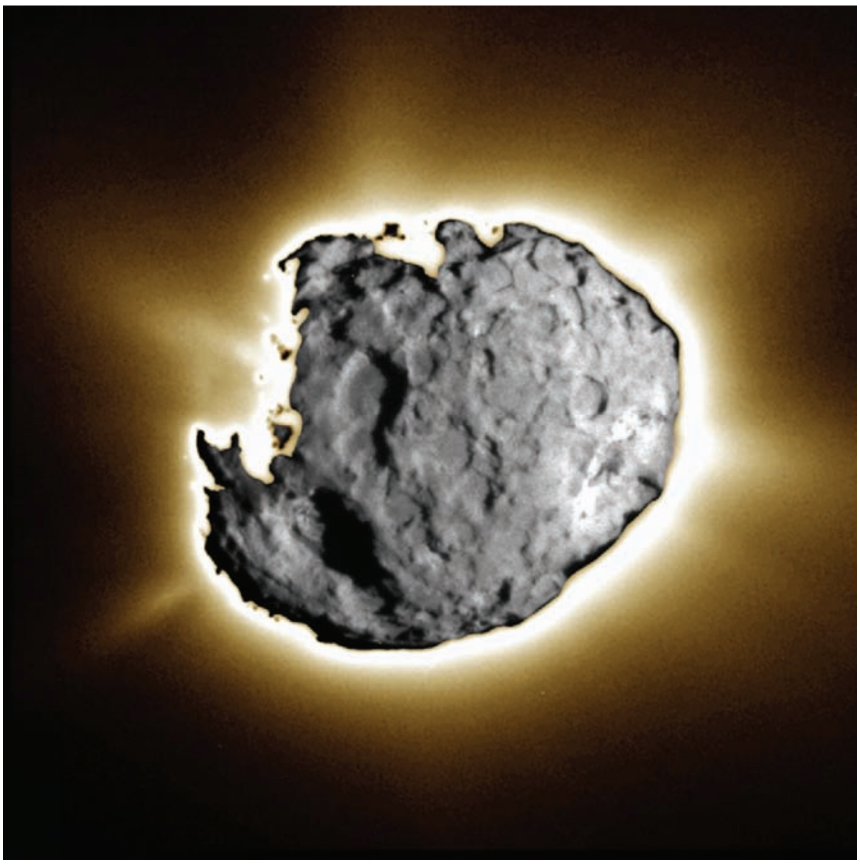


Fig. 14.17 Nucleus of comet Wild 2 An image of the dark, cratered nucleus of comet Wild 2 taken from the *Stardust* spacecraft during its comet flyby on 2 January 2004. The nucleus is about 5 kilometers in diameter. This is a composite of a short exposure image showing surface details, and a long exposure image, taken just 10 seconds later, showing the active surface jetting dust and gas streams into space. During this flyby *Stardust* gathered comet dust and subsequently returned the comet sample to Earth. (Courtesy of NAS/JPL-Caltech.)

Focus 14.3 High-temperature and organic materials found in comet dust return

Comet particles returned from comet Wild 2 by the *Stardust* spacecraft originated in the frozen nucleus of the comet, which has hibernated for billions of years in one of the coldest places in the solar system and has not been significantly altered by repeated passages near the Sun. Yet the returned comet dust included high-temperature minerals, and contained a mix of material that must have formed at some of the highest and lowest temperatures that existed in the early solar system. These discoveries might alter views of comet formation.

Two kinds of nitrogen-rich organic molecules and the amino acid glycine have also been discovered in samples of comet Wild 2. These results indicate that comets could have delivered similar organic compounds to the Earth along with water, especially during the planet's early history when comet impacts were more frequent than now. The rain of comet dust may have even brought the self-replicating, carbon-rich molecules necessary for the origin of life, but this remains a speculation without hard scientific evidence.

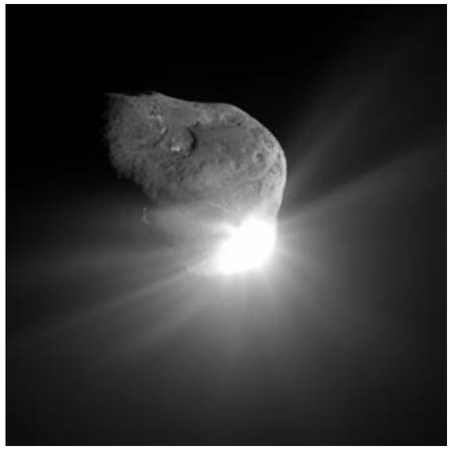


Fig. 14.18 Impact with nucleus of comet Tempel 1 The initial material ejected when the *Deep Impact probe* collided with the nucleus of comet Tempel 1 on 4 July 2005. The dark, cratered nucleus is an oblong body between 4.9 and 7.6 kilometers in size. Spectroscopic examination of the ejected material showed that the comet nucleus is mainly composed of water ice and dust. (Courtesy of NASA/JPL-Caltech/UMD.)

would image the comet, and a smaller impactor that would collide with the comet and find out what is inside it. The collision occurred as scheduled on 4 July 2005, in celebration of the American Independence Day.

Moving at a speed of 10 kilometers per second, the 820-pound impactor vaporized deep below the comet's surface and kicked up a cloud of dust (Fig. 14.18). The opacity of the plume that the impactor created, and the light it gave off, suggested that the dust excavated from comet Tempel 1 is very fine and porous, like snow rather than sand. It might cover an interior that has remained unchanged since the comet formed 4.6 billion years ago.

Telescopes trained on the ejecta detected spectra of water vaporized by the heat of impact, but also the absorption spectra of ice particles ejected from the surface a few seconds after the water vapor was initially released. This was the first time direct evidence had been obtained for water ice in a comet and on its surface – not melted, vaporized or sublimated but solid ice. Most of this escaping water was contained in ice particles from below the surface.

Instruments aboard the *Spitzer Space Telescope* detected a huge increase in the amount of carbon-containing molecules, including carbon dioxide and hydrocarbons, detected in the spectral analysis of the ejection plume, indicating that comets contain substantial amounts of organic material.

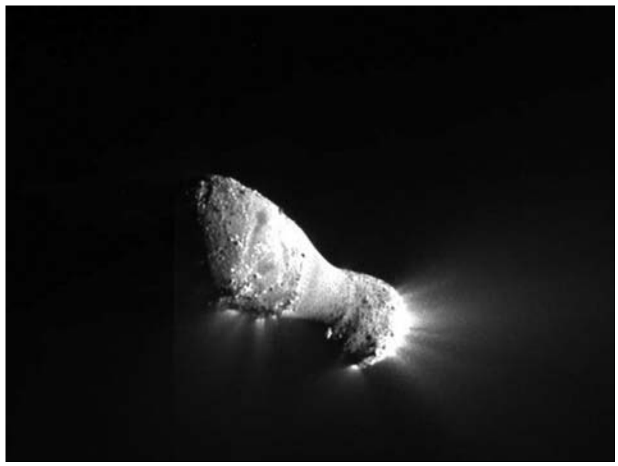


Fig. 14.19 The nucleus of comet Hartley 2 Jets stream out of the sunlit side (*right*) of the nucleus of the small, active comet Hartley 2, whose long axis spans about 2 kilometers. A narrower, smooth waist joins rough areas at both ends of the elongated nucleus. (Courtesy of NASA/JPL-Caltech/UMD.)

After releasing its impact probe, the *Deep Impact* spacecraft was retargeted to fly past comet Hartley 2, as part of the extended *EPOXI* mission. The closest approach, at a distance of 700 kilometers, revealed a small, oblong nucleus with activity on the sunlit side and two rough areas joined at a smoother middle (Fig. 14.19).

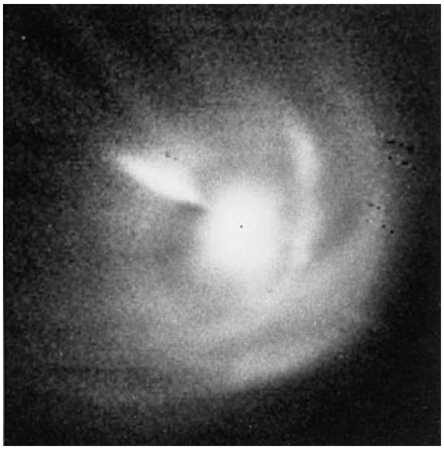


Fig. 14.20 Rotating comet This photograph of comet Halley shows jets of dust ejected from a rotating nucleus. (Courtesy of Stephen Larson, LPL/U. Arizona.)

A *Rosetta* future

On 2 March 2004, the European Space Agency launched the *Rosetta* spacecraft on a complex 10-year journey to comet Churyumov-Gerasimenko, using four gravity assists, three from Earth and one from Mars. Beginning in August 2014 the spacecraft is expected to spend approximately 2 years orbiting and observing the comet as it approaches the Sun. *Rosetta* is also expected to place a small lander on the comet's surface in November 2014.

14.7 Rotating comet nucleus

A comet's nucleus also rotates (Fig. 14.20). Typical rotation periods are a few hours to a few days. Observations of comet Halley, for example, indicate that it rotates around its longest axis once every 7.4 days, and that it wobbles about its shortest axis once every 2.2 days, or 53 hours. As the nucleus rotates, new regions turn to face the Sun, heat up and become active, while others face away from the Sun and momentarily turn off their activity.

Jets from a rotating comet nucleus can help explain comets that seem to defy gravity by arriving at perihelion a few days before or after the time calculated using Newton's theory of gravitation. The gas and dust streaming off the sunlit side of a comet nucleus initially heads toward the Sun, before being swept back into the comet tails. The expelled material pushes the comet in the opposite direction, making it arrive sooner or later than expected (Fig. 14.21). A similar recoil effect explains the darting action of a small balloon when it is released, as well as the forward thrust of a rocket engine. The thrust of the rocket-like jets either pushes the comet along in its orbit or slows it down.

14.8 Comet decay and meteor showers

Comets fall apart

Once a Jupiter-family comet enters the inner solar system, it returns again and again on a relentless voyage of continual disintegration, eventually being consumed by their own emissions. They are caught in a life of continual decay, sublimating and blowing part of themselves away each time they come near the Sun.

Sooner or later most Halley-type and Jupiter-family comets will either fall apart or turn into a dark rocky corpse that looks like an asteroid. It's just part of the aging process. Once you can see one of these comets, they are doomed to disappear.

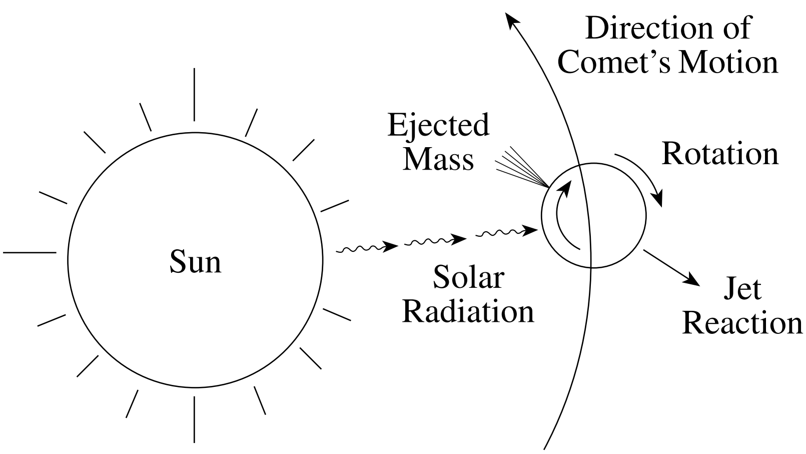


Fig. 14.21 How to make a jet engine out of a dirty ball of comet ice Unexpected cometary motions are attributed to non-gravitational forces caused by jets of matter ejected from a comet's spinning, icy nucleus. In this illustration, the ejected material pushes the comet in the opposite direction to its motion, causing the comet to arrive closest to the Sun at a time later than expected. If the comet had been rotating in the opposite direction, the jets would have pushed the comet along in its original direction, resulting in an early arrival time.

Fig. 14.22 Comet clones The splitting of the nucleus of comet West (1976 VI) photographed (*top to bottom*) on 8, 12, 14, 18 and 24 March 1976, in yellow-green light using a 0.60-meter (23.6-inch) Cassegrain reflector. On 18 March, the diameter of the four features was about 10 000 kilometers. (Courtesy of C. Knuckles and S. Murrell, New Mexico State University Observatory.)

Comets are very fragile, with little internal strength and a very low mass density. Some comet nuclei are crumbly, fluffy structures with mass densities less than that of solid ice and far less than solid rock. The central pressure of their nucleus is probably comparable to that under a thick layer of blankets. So it is little wonder that some comets have been observed to break up as the result of tidal forces from either the Sun or Jupiter. They pull on the near side of the comet a little more than the far side, tearing the comet apart. The nucleus of comet West was, for example, split into pieces when it passed near the Sun in 1976 (Fig. 14.22).

In other instances, a comet has spontaneously split apart, with no apparent gravitational or tidal disruption from Jupiter or the Sun. Sometimes a comet just blows off its fragile outer parts with no apparent reason (Fig. 14.23).

The pieces of the split nucleus have too little mass to pull themselves together gravitationally. So once a nucleus splits, its pieces remain forever separated. The jets of escaping gas kick them away from each other, and they continue to drift farther apart.

Any Jupiter-family comet is slowly wasting away from the outside in. Each time it approaches the Sun, a small percentage of a comet's icy surface will sublimate and escape into space, dragging along grains of dust. An example is comet Encke that revolves about the Sun with a 3.3-year orbital period; it loses about a meter of ice and dust each time it passes near the Sun. If its nucleus were 3 kilometers across, then this comet would disappear in just 10 000 years. Comet Halley is not expected to last more than about 76 000 years, and most of the comets seen during recorded history will probably vanish from sight in less than a million years, either sublimating away into nothing or leaving behind a black, rocky, burned-out corpse.

Nights of the shooting stars

Spectacular meteor showers, spawned by passing comets, have periodically returned for at least a thousand years, inspiring fear, wonder and admiration. William Blake (1757–1827) caught some of the excitement of the August shower, the Perseids, in his poem "The Tyger":

> When the stars threw down their spears
> And water'd heaven with their tears . . .

Meteor showers are commonly called shooting stars (Fig. 14.24), but they are not stars. They are fragile material from comets. In addition to spewing off small dust particles, which can be captured by Earth, drift down to the ground, and enter your hair, comets also expel larger particles ranging in size from sand grains to pebbles. This debris burns up when it enters our atmosphere, producing visible meteors.

The meteors associated with comets vaporize completely in flight. In fact, no meteor associated with a comet has ever reached the ground. As a cometary meteoroid enters the Earth's atmosphere, it will decelerate and heat up as the result of friction, producing a luminous trail until the object becomes hot enough to vaporize completely away. Thus, there is a great difference between the fragile meteoric material associated with comets and the tough meteorites that survive the atmospheric flight to the ground.

When just one of the comet particles rubs against the air, it vaporizes in a streak of light, producing the luminous

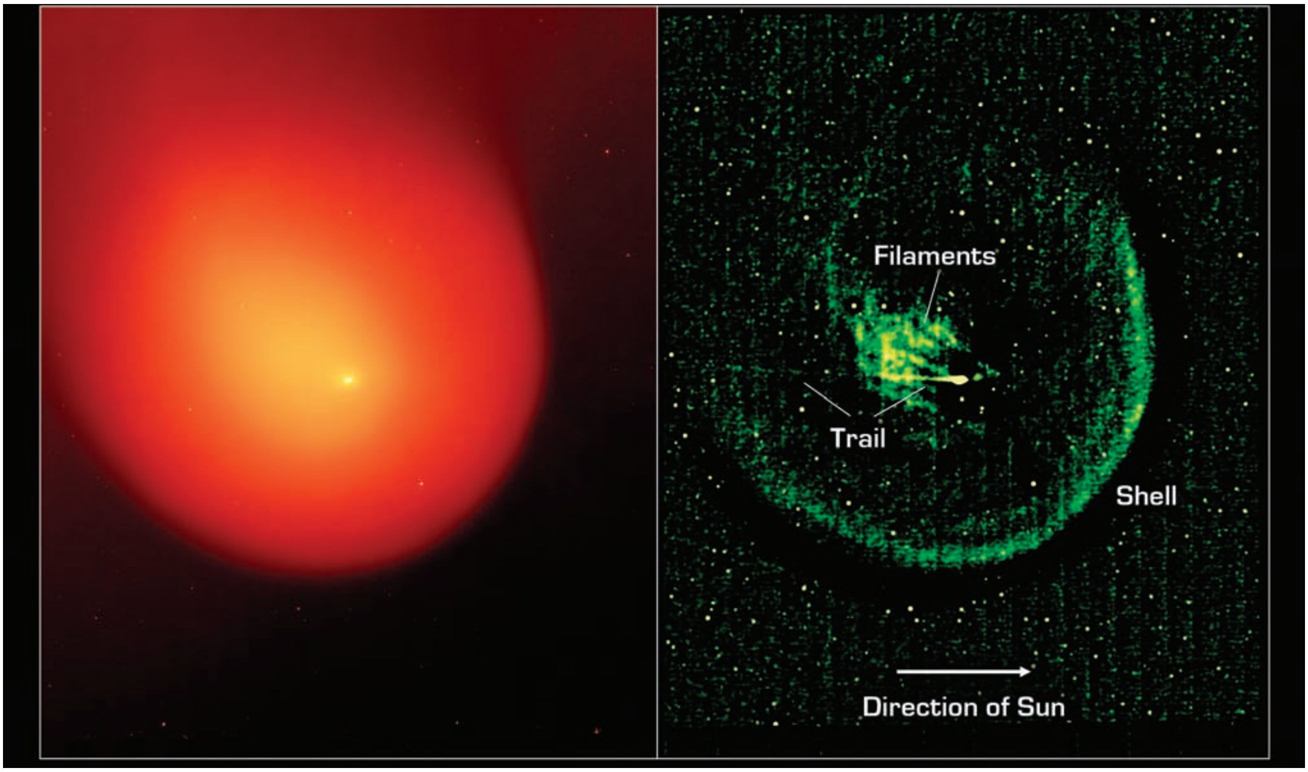


Fig. 14.23 Comet explosion Comet Holmes has exploded just twice in 118 years, in November 1892 and October 2007, while approaching the asteroid belt and moving from the vicinity of Jupiter toward the Sun. This infrared image was obtained from the *Spitzer Space Telescope* in March 2008; five months after the comet suddenly erupted and brightened a million-fold overnight. The infrared picture (*left*) reveals dust particles that make up the outer shell, which envelops solid particles (*yellow*) blown from the exploding comet and the comet nucleus (*white center*). The contrast-enhanced image (*right*) shows the outer shell and filaments or streamers of dust. (Courtesy of NSA/JPL-Caltech.)

Fig. 14.24 The shooting stars Two couples portray meteor showers or falling stars in this picture painted by Jean-Francois Millet (1814–1875) in 1847. They soar through the skies, perhaps illustrating the transcendental nature of erotic love. (Courtesy of the National Museum of Wales, Cardiff.)

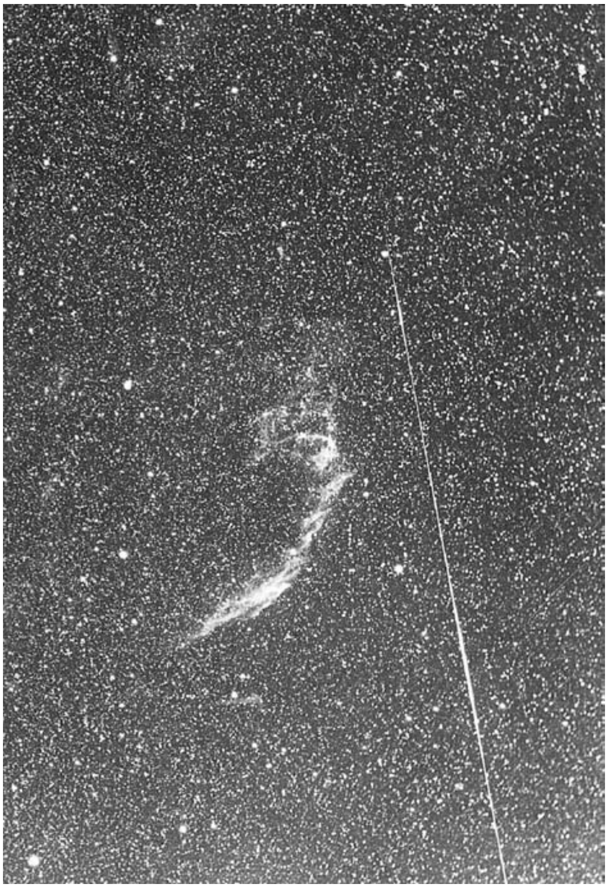


Fig. 14.25 Meteor trail A glowing meteor streaks across the stars near the constellation Cygnus. The straight trail was produced by a sand- or pebble-sized piece of a comet, burned up by friction as it entered the Earth's atmosphere. The curved structure (*left center*), known as the Cygnus Loop, is an expanding shell of material thrown off during the supernova explosion of a massive, dying star. (Courtesy of the Yerkes Observatory.)

Table 14.6 Comets associated with meteor showers[a]

Meteor shower[b]	Comet
Lyrids	Thatcher C/1861 G1
Eta Aquarids	1P/Halley
Scorpiids-Sagittariids	2P/Encke
Bootids	7P/Pons-Winnecke
Perseids	109P/Swift-Tuttle
Aurigids	Kiess C/1911 N1
Draconids	22P Giacobini-Zinner
Orionids	1P/Halley
Epsilon Geminids	Ikeya C/1964 N1
Taurids	2P/Encke
Adromedids	3P/Biela
Leonids	55P/Temple-Tuttle
Geminids	3200 Phaëthon[c]
Monocerotids	Mellish C/1917 F1
Ursids	8P/Tuttle

[a] Adapted from Kenneth R. Lang, *Astrophysical Data: Planets and Stars*, New York, Springer Verlag (1992).
[b] The visibility dates of these showers are given in Table 14.7.
[c] 3200 Phaëthon is cataloged as an asteroid, but it may be an inactive comet nucleus.

trail of a meteor (Fig. 14.25). And when many fall into the dark night sky, they produce meteor showers.

From the luminous path of a meteor, it is possible to determine the incoming particle's orbital path around the Sun, and in most cases the orbits are similar to those of comets (Table 14.6). A comet ejects the particles along its orbital path as it loops around the Sun, and this material continues to revolve around our star. The swarm of comet material is called a meteoroid stream, which is spread out all along the orbit of the comet (Fig. 14.26). And when the Earth passes though one of these streams, it intercepts some of the orbiting particles that enter our atmosphere, and creates a meteor shower (Fig. 14.27).

When a meteor shower includes large numbers of shooting stars, the trails appear to intersect and emanate from a distant point called the *radiant* (Fig. 14.28). But meteors that appear to diverge from a point are actually moving on parallel paths, just as parallel railroad tracks seem to come from a point on the distant horizon. Meteor showers are named after the constellation in which their radiant appears (Table 14.7).

Because the Earth often passes through a comet's orbit just once a year, a meteor shower usually appears at yearly intervals. The Lyrids, for instance, appear in April, the Perseids are seen every August, and the Leonids light up the night sky in November. Since the distribution of material along a comet orbit is generally non-uniform, the rates of meteors in a particular shower can vary from year to year.

The Earth has been passing through some comet orbits for hundreds of years. The Lyrids were known in 687 BC, the Perseid stream was first recorded in 36 AD and the Leonids were first recorded in 902 AD. Nevertheless, the comet that produces a meteor shower will eventually disintegrate completely, and that shower will not be seen again. But this may not happen for a million years from now.

Burned-out comets that look like asteroids

Although comets are distinguished by their ability to emit gas and dust when near the Sun, some comets may be either dormant or extinct. In fact, there are a few small

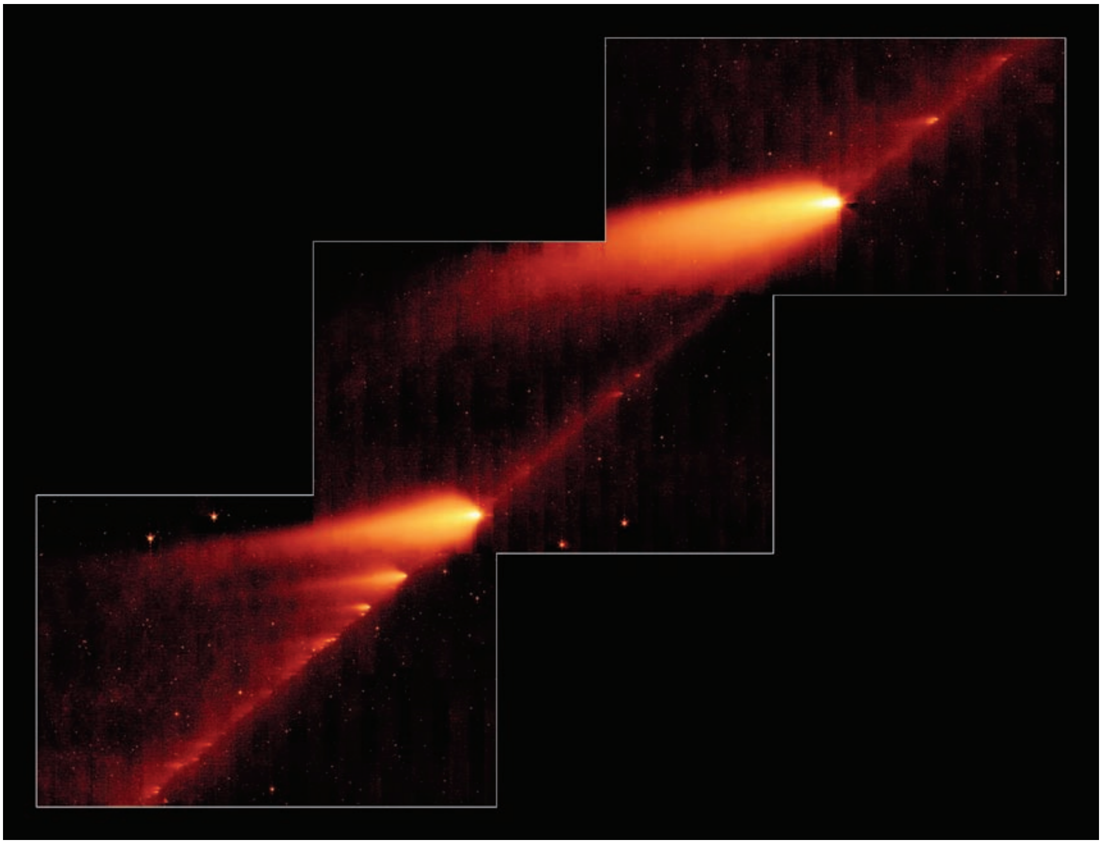


Fig. 14.26 Meteoroid stream from disintegrating comet The broken comet Schwassman-Wachmann 3 moves along a trail of debris formed during its multiple trips around the Sun, which the comet circles every 5.4 years. In 1995 the comet broke apart, and two of the biggest fragments are shown here, as miniature comets, together with about 36 smaller fragments. The line bridging the large comet fragments is the meteor stream of dust, pebbles and rocks left in the comet's wake during numerous previous journeys. This infrared image was taken on 4 to 6 May 2006 from the *Spitzer Space Telescope*. (Courtesy of NASA/JPL-Caltech.)

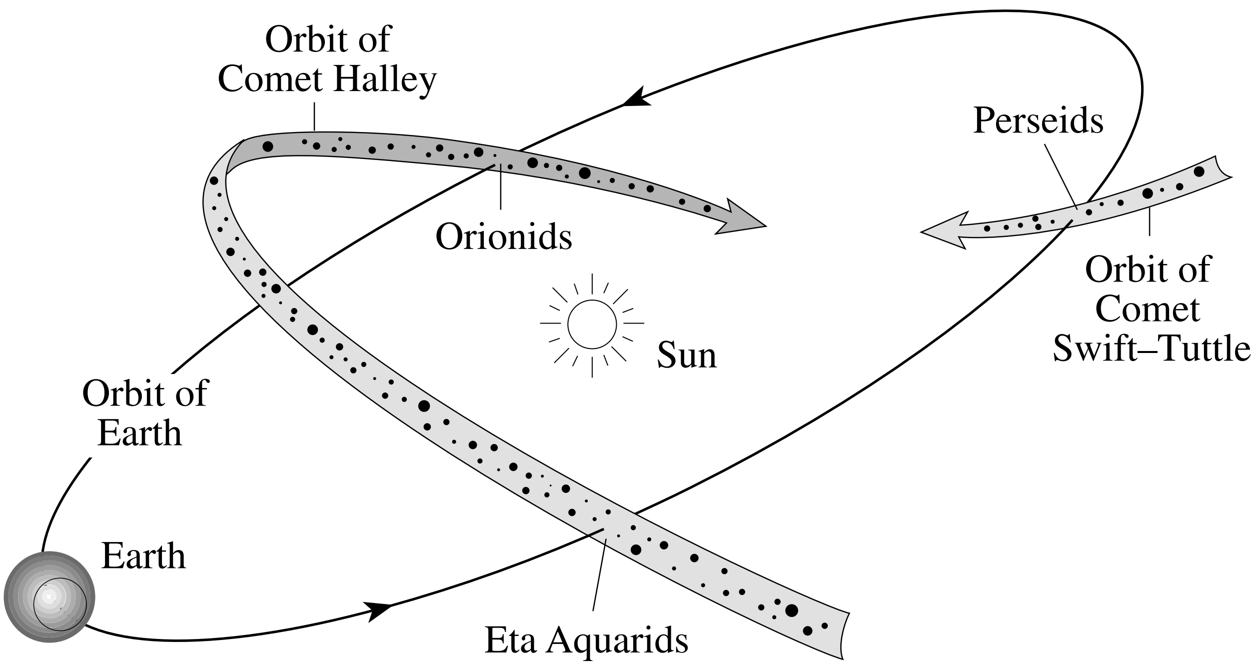


Fig. 14.27 Comets produce meteor showers The Earth's orbit intersects a stream of meteoric material left along the orbit of comet Halley, producing two meteor showers, the Eta Aquarids in May and the Orionids in October. Other comets intersect the Earth's orbit just once during their trip around the Sun. Annual meteor showers are created when the Earth enters the intersection point, such as the August Perseids produced by debris from comet Swift–Tuttle. The orbit of comet Halley is inclined by 162 degrees with respect to the ecliptic, the plane of the Earth's orbit, while the orbit of comet Swift–Tuttle has an inclination of 114 degrees.

Table 14.7 The principal annual nighttime meteor showers[a]

Shower	Maximum date	Radiant position: Right ascension	Radiant position: Declination	Visibility dates	Meteors per hour
Quadrantids	3–4 January	15h 28m	+50°	1–6 January	110
Alpha Aurigids	6–9 February	04h 56m	+43°	Jan.–Feb.	10
Virginids	12 April	14h 04m	–09°	March–April	5
		13h 36m	−11°		5
Lyrids	21–22 April	18h 08m	+32°	19–25 April	15
Eta Aquarids	5 May	22h 20m	−01°	24 April–20 May	35
Alpha Scorpiids	28 April	16h 32m	−24°	20 April–19 May	5
	13 May	16h 04m	−24°		
Ophiuchids	9 June	17h 56m	−23°	May–June	5
	19 June	17h 20m	−20°		
Alpha Cygnids	21 July	21h 00m	+48°	June–August	5
Capricornids	8–15 July	20h 44m	−15°	5 July–20 August	5
Alpha Capricornids	2 August	20h 36m	−10°	15 July–20 August	5
Delta Aquarids	29 July	22h 36m	−17°	15 July–20 August	25
	6 August	23h 04m	+02°		10
Iota Aquarids	6 August	22h 10m	−15°	July–August	10
Piscis Australids	31 July	22h 40m	−30°	July–August	5
Perseids	12 August	03h 04m	+58°	25 July–20 August	80
Alpha Aurigids	28 August	04h 56m	+43°	August–October	10
Piscids	9 September	00h 36m	+07°	Sept.–Oct.	10
Orionids	21 October	06h 24m	+15°	15 Oct.–2 Nov.	30
Taurids	3 November	03h 44m	+14°	15 Oct.–25 Nov.	10
Leonids	17 November	10h 08m	+22°	15–20 November	45
Geminids	13–14 December	07h 28m	+32°	7–15 December	70
Ursids	22–23 December	14h 28m	+78°	19–24 December	10

[a] The celestial coordinates of the radiant are right ascension (RA), in hours (h) and minutes (m), and declination (dec.), in degrees. The maximum hourly frequency of meteors assumes the radiant is at the zenith, but this rate can vary from year to year because of the non-uniform distribution of the relevant meteoroid stream. [Adapted from N. Bone, *Meteors*, Cambridge, MA, Sky Publishing Co. (1993), and Kenneth R. Lang, *Astrophysical Data: Planets and Stars*, New York, Springer Verlag (1992).]

objects that move like comets, but emit no gas and dust and display neither a coma nor a tail. Some of them have eccentric, elongated comet-like orbits that stretch into the vast outer reaches of the solar system, beyond the most distant major planets. Yet they show no trace of comet activity when approaching the warmth of the Sun, sometimes passing nearer to it than the Earth.

They may be inert comets, which have turned into inactive objects. These comets either exhausted all the volatile ices they once had to feed a coma and tail, or their ice might be completely shrouded in a thick, insulating cover of dust and dirt, preventing the ice from sublimating into luminous material.

The near-Earth asteroid 3200 Phaëthon is a well-known example. It moves in a highly eccentric orbit – once thought to be the hallmark of comets. In fact, it follows the orbit of the meteoroid stream that produces the Geminids meteor shower. Since the large majority of meteor showers are caused by debris scattered along a comet's orbit, it is likely that this object is a defunct comet that has now lost its ability to emit gas and dust.

These anomalous bodies account for a very modest fraction of the thousands of known comets and asteroids.

We now turn our full attention to other larger objects in the Kuiper belt beyond Neptune.

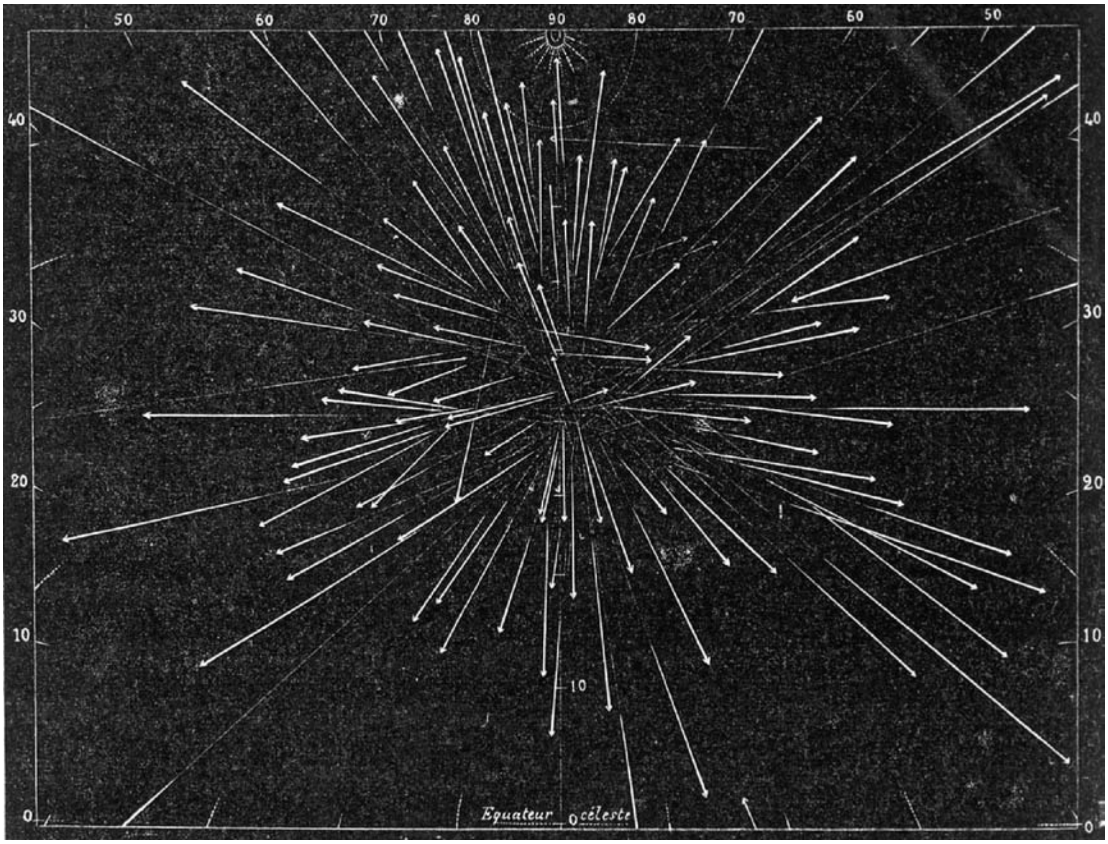


Fig. 14.28 Radiant meteors The apparent paths of shooting stars on 27 November 1872. Meteor showers are named after the constellation in which their radiant appears. This meteor shower is called the Andromedids meteor shower because its radiant appears in the constellation Andromeda. The shower occurs every November when the Earth intersects the debris that has been scattered along the orbit of Biela's comet. [Adapted from Amedee Guillemin's *Le Ciel*, Librairie Hachette, Paris (1877).]

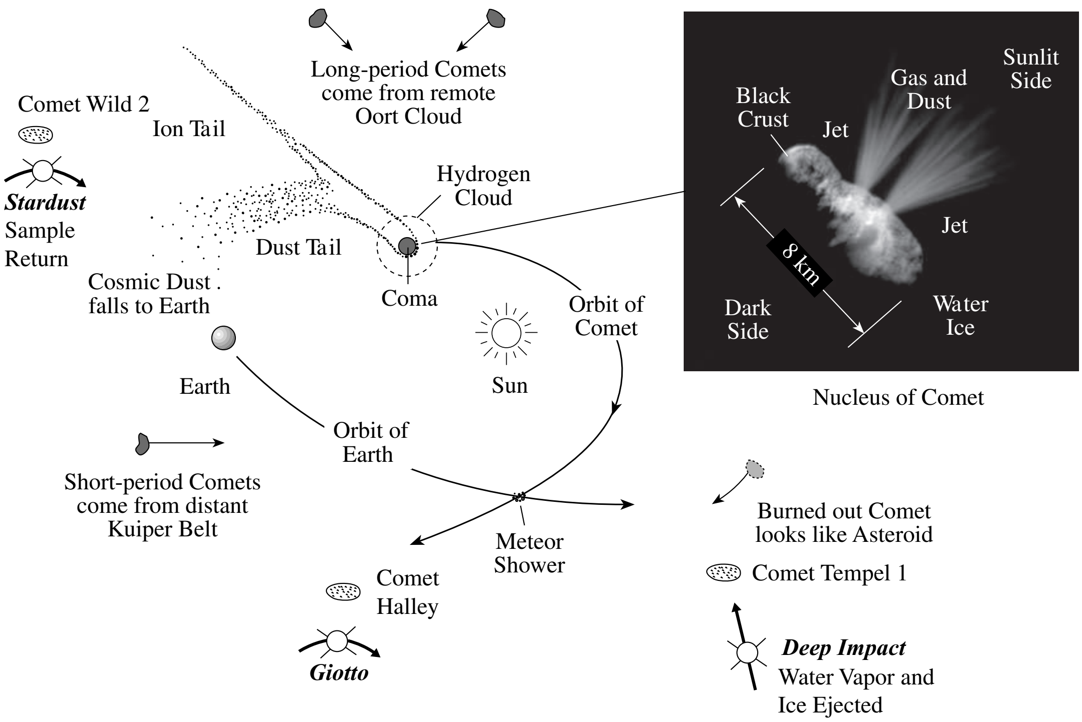


Fig. 14.29 Summary diagram

15 Beyond Neptune

- The discovery of Pluto was serendipitous, the result of a long careful search based on incorrect predictions.
- Pluto moves between 29.7 and 49.3 AU from the Sun, and just crosses within the orbit of Neptune, at 30 AU, without ever colliding with it.
- Pluto has a thin atmosphere of nitrogen, methane and carbon dioxide, and large variable dark and light surface makings.
- Pluto has three companion moons, named Charon, Nix and Hydra. Charon is an oversized satellite, about half the size of Pluto.
- The *New Horizons* spacecraft will encounter Pluto in 2015, to study its atmosphere, surface and interior, then travel into the Kuiper belt.
- Centaurs are small bodies that orbit the Sun between Jupiter and Neptune and cross the orbit of one or more giant planets. They are thought to originate further away from the Sun in the Kuiper belt. Some centaurs display comet comas, and most of them are expected to become comets.
- Several trans-Neptunian objects orbit the Sun beyond the orbit of Neptune, including Eris, Makemake and Haumea. They are either larger than Pluto or comparable to it in size.
- Pluto has been demoted from the ninth major planet to a dwarf planet, and then reclassified as a plutoid.
- A dwarf planet orbits the Sun and has a rounded shape, but it does not clear out the orbit around it. The asteroid 1 Ceres has been designated a dwarf planet.
- Plutoids are trans-Neptunian dwarf planets, and they include Eris, Haumea, Makemake and Pluto.
- The *Voyager 1* and *2* spacecraft have crossed the termination shock of the solar wind at 94 AU and 84 AU respectively.
- The outer Oort cloud lies between 20 000 and 200 000 AU, far beyond the termination of the solar wind, but still within the Sun's gravitational control.

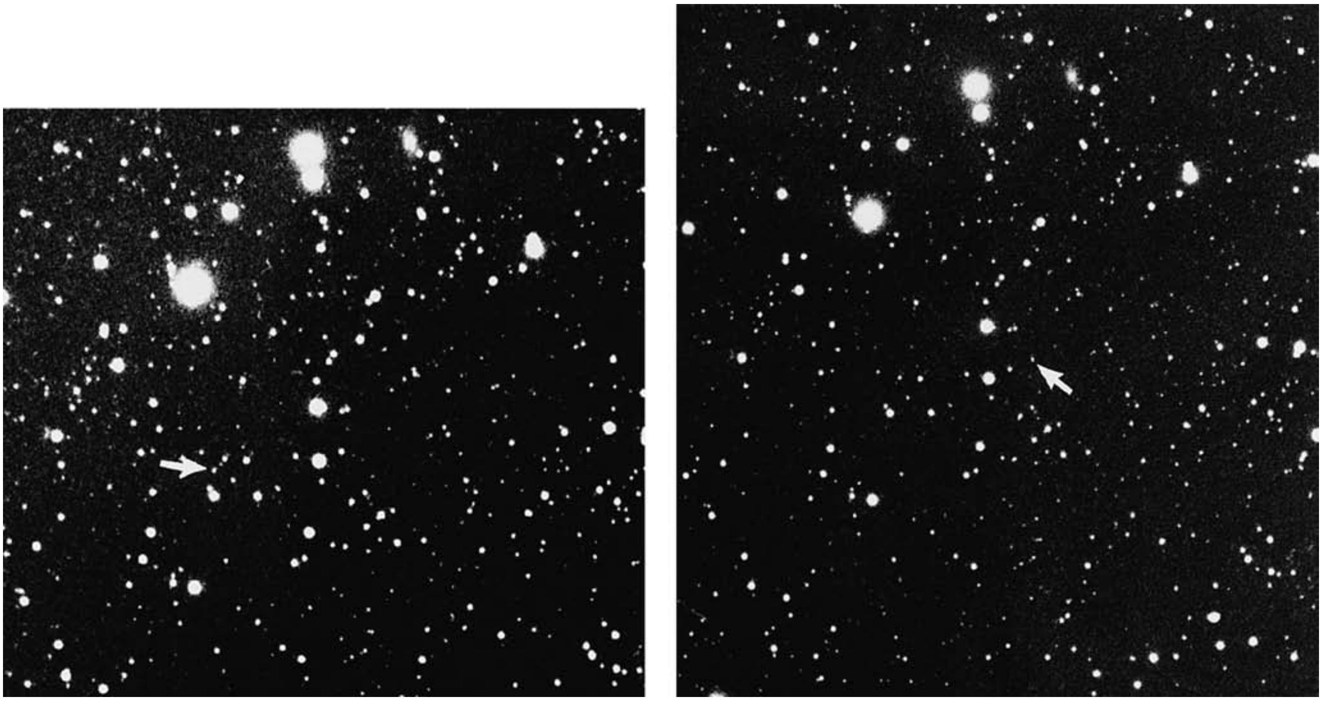


Fig. 15.1 Discovery of Pluto A region of the constellation Gemini, photographed by Clyde W. Tombaugh (1906–1997) on 23 January 1930 (*left*), and the same region photographed six days later (*right*). When comparing the two plates on 18 February 1930 with a blink microscope, Tombaugh noticed an object (*arrows*) on the second plate that had changed its location with respect to the background stars since the first plate was taken. This was a previously unknown object that had to belong to the solar system. Because of its slow apparent motion across the sky, the planet images were separated by just 3.5 millimeters on the two photographs. (Courtesy of the Lowell Observatory.)

15.1 Pluto: a small frozen world with companions

The serendipitous discovery of Pluto

The discovery of Neptune in 1846 resulted from a mathematical study of the differences between the predicted and observed positions of Uranus, attributed to the gravitational pull of the then unknown planet. Astronomers hoped that similar irregularities in Neptune's motion would lead to the discovery of another remote planet; but because of Neptune's long 165-year orbit there were insufficient observations. Prediction of another unknown planet therefore had to be based upon perturbations in Uranus' motion, after corrections for the gravitational effects of Neptune.

Two astronomers used the corrected Uranus data to predict an undiscovered planet beyond Neptune. The first such prediction was made in 1909 when William Henry Pickering (1858–1938) argued that both Neptune and a remote Planet O were producing gravitational tugs on Uranus. Percival Lowell (1855–1916) made the next attempt in 1915; he called his unknown object Planet X.

Lowell directed the most ambitious search for the trans-Neptunian planet, at his observatory in Flagstaff, Arizona, but no new planet was found at a variety of predicted locations between 1905 and 1919. The search from the Lowell Observatory continued a decade later using a new 0.33-meter (13-inch) photographic telescope. Once three photographs had been taken at intervals of several days, they were set in pairs in a blink microscope that would show the apparent motion of a planet, asteroid or comet against a background of nearly half a million stars on each photograph.

After months of painstaking work, Clyde William Tombaugh (1906–1997) discovered, on 18 February 1930, the sharp, faint, moving image of the elusive quarry (Fig. 15.1). The new object was named Pluto, for the Roman god of the underworld. It is a small frozen world at the outer fringe of the planetary system, traveling along an elongated, 248-year path that is inclined by 17 degrees to the orbital plane of the major planets.

Pluto's orbit around the Sun is so far from circular that it moves between 29.7 and 49.3 AU from the Sun, and it crosses the orbit of Neptune at 30.06 AU. Pluto is, however, protected from possible collision with the large planet by a special orbital relationship, called a

Table 15.1 Physical characteristics of Pluto

Mean radius	1153 kilometers
Mass	1.309×10^{22} kilograms
Bulk density	2050 ± 40 kilograms per cubic meter
Sidereal rotation period	−6.3872 Earth days
Sidereal orbital period	247.92 years
Average distance from Sun	39.48 AU

resonance. Simply put, for every three orbits of Neptune around the Sun, Pluto completes two. This means that although Pluto's orbital plane crosses that of Neptune, the two bodies never come closer than 17 AU from each other.

Pluto is an anomaly. It is much smaller than the giant planets, and is comparable in size to some of their satellites (Table 15.1). Pluto is smaller than Saturn's satellite Titan and all four of Jupiter's largest moons. In many ways, Pluto is the twin of Neptune's largest satellite Triton. They have almost the same size and mean mass density. Moreover, many other small worlds have now been discovered just beyond the orbit of Neptune. Pluto is more akin to this group of objects than to the families of terrestrial or giant planets.

The discovery of an object that orbits Pluto led to a determination of its mass, at about only 1.3×10^{22} kilograms, or just 0.2 percent (0.002) of the Earth's mass and only about one-sixth the mass of the Earth's Moon. This means that Pluto was not found because it was correctly predicted. Its mass is far too small to have noticeably influenced the past motions of Uranus. In fact, the distant Earth exerts a larger gravitational influence on Uranus than Pluto does. The discovery of Pluto was the result of a meticulous and systematic search that was guided by an incorrect prediction, which merely happened to point in the general direction of Pluto.

After Pluto's discovery, astronomers continued to speculate that some unknown, massive and remote planet was responsible for the apparent perturbations in the motion of Uranus. However, after accounting for the gravitational effects of Neptune, using a precise mass obtained when *Voyager 2* encountered the planet in 1989, the small unexplained differences between the predicted and observed locations of Uranus simply disappeared. This means that there is no massive trans-Neptunian planet, and that all of the sizeable planets in our solar system have been discovered. Neptune is effectively the outermost major planet, serving as a lonely distant sentinel to outer space.

Pluto's atmosphere and surface

In 1998 Pluto passed in front of a star, revealing a thin, extended atmosphere that caused a gradual reduction in the star's light. If there were no atmosphere, the starlight would vanish abruptly. But Pluto does not have much of an atmosphere, for its surface pressure is about one-millionth that of Earth's.

The atmosphere's three main gases – nitrogen, methane and carbon monoxide – will partially freeze onto Pluto's surface during the 100-year long period when it is farthest from the Sun. When closest to the Sun, the gases will slowly escape into space because of Pluto's low gravity.

Spectroscopic evidence, first obtained in 1976, shows that methane frost covers much of the surface of Pluto. Nitrogen frost has also been detected on Pluto. A comparison of *Hubble Space Telescope* (*HST*) images of Pluto taken in 1994 with those taken in 2002–03 indicates that its surface is gradually changing with the seasons. *HST* images also show that large-scale dark and light markings are distributed around Pluto's globe (Fig. 15.2). They may be attributed to topographic features or to the distribution of frosts.

Pluto's companions

Pluto has an oversized companion that is half as big as Pluto is. This discovery was an accidental by-product of observations made for another purpose. In 1978, astronomers at the United States Naval Observatory were obtaining a series of photographs to improve the accuracy of Pluto's orbit, when several of the images appeared slightly distorted, from a round to oblong shape. The elongation seemed to disappear every few days, and careful examination showed that it is caused by another small world that orbits Pluto. The two objects are so close to each other and so far away that they remain blurred together when viewed with even the best telescopes on the ground, but they can be clearly resolved with the *Hubble Space Telescope* that orbits the Earth above its obscuring atmosphere (Fig. 15.3).

Pluto's companion is named Charon, after the boatman who ferried new arrivals across the river Styx at the entrance to Pluto's underworld, Hades. Penniless ghosts are said to have waited endlessly because Charon gave no free rides.

The announcement of this remarkable doubling was a happy surprise, for it permitted determining the mass of Pluto. Charon orbits Pluto at a distance of 19 640 kilometers once every 6.387 Earth days. For comparison, if our Moon were that close to Earth it would orbit in 7 hours.

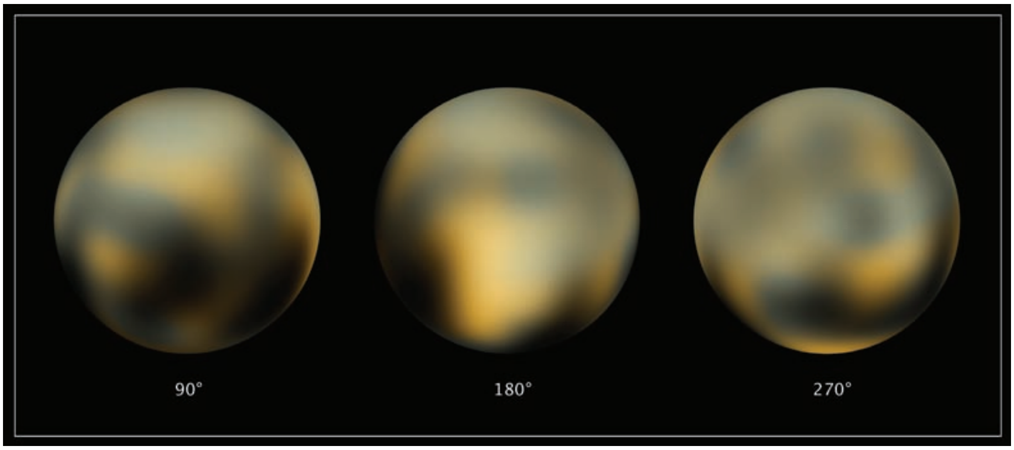


Fig. 15.2 Pluto's changing surface Dark-orange and charcoal-black terrain is seen on three sides of Pluto. Ultraviolet radiation from the Sun is thought to break up methane that is on the dwarf planet's surface, leaving behind a dark carbon-rich residue. The center image, at 180 degrees, has a bright spot that is unusually rich in carbon-monoxide frost. These images were constructed from multiple *Hubble Space Telescope* (*HST*) observations in 2002–03. When compared with *HST* images taken in 1994, seasonal changes in color and brightness are detected, probably created when ices melt and refreeze and the tenuous atmosphere changes. Although these views are not sharp enough to resolve mountains or craters, they provide background for closer scrutiny when the *New Horizons* spacecraft encounters Pluto in 2015. (Courtesy of NASA/ESA/Marc Buie, SRI.)

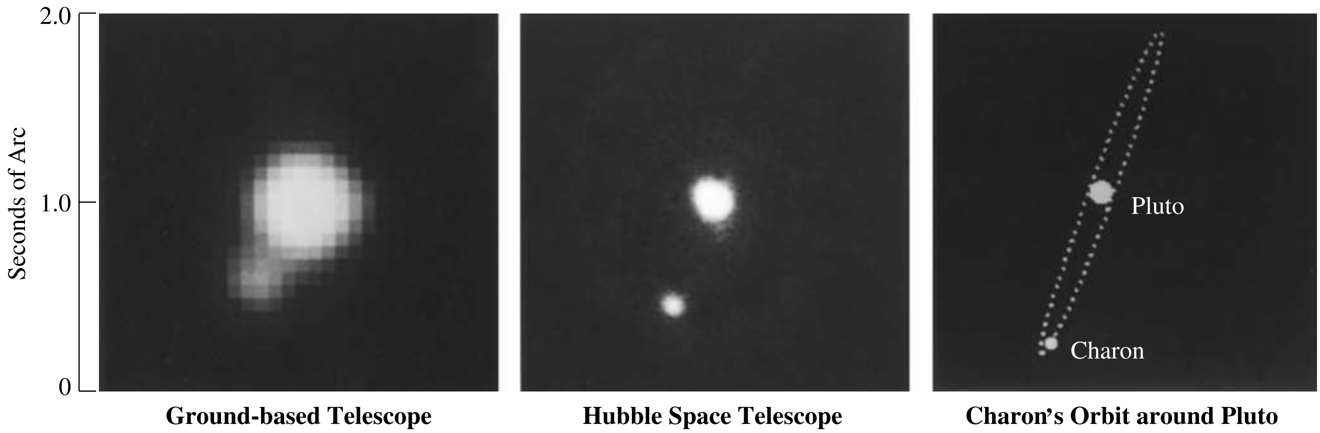


Fig. 15.3 A double object The *Hubble Space Telescope* (*HST*) distinguishes between Pluto, the bright object at the middle of the central image, and its companion Charon, which is the fainter object in the lower left of the central image. Observations from telescopes on Earth (*left*) were unable to clearly resolve the pair because of atmospheric distortions. At the time of the *HST* photograph, Charon's orbit around Pluto (*right*) was seen nearly edge-on, and Charon was near its maximum angular separation from Pluto, a mere 0.9 seconds of arc. Pluto's mean radius is 1153 kilometers, Charon is half that size, and the two objects are just 19 640 kilometers apart. The *HST*'s ability to distinguish Pluto's disk at its distance of 4.4 billion kilometers is equivalent to seeing a baseball at a distance of about 100 kilometers. (Courtesy of NASA.)

Charon's slow revolution about Pluto is a result of Pluto's small mass.

When Pluto and Charon pirouette into an edge-on view from Earth, we see a series of mutual eclipses as the two objects alternately pass directly in front of each other. Timing the starts and ends of such occultations permit an accurate determination of their size. Pluto has a mean radius of just 1153 kilometers, or about two-thirds that of the Earth's Moon. Charon is nearly half the size of the Pluto, with a mean radius of 603.5 kilometers.

In mid-2005, almost three decades after Charon's discovery, a team of astronomers found two more moons of Pluto as the result of a dedicated search with the *HST* (Fig. 15.4). The new moons have been named Nix, for the mother of Charon and the goddess of darkness, and Hydra, for the nine-headed serpent of the underworld. The

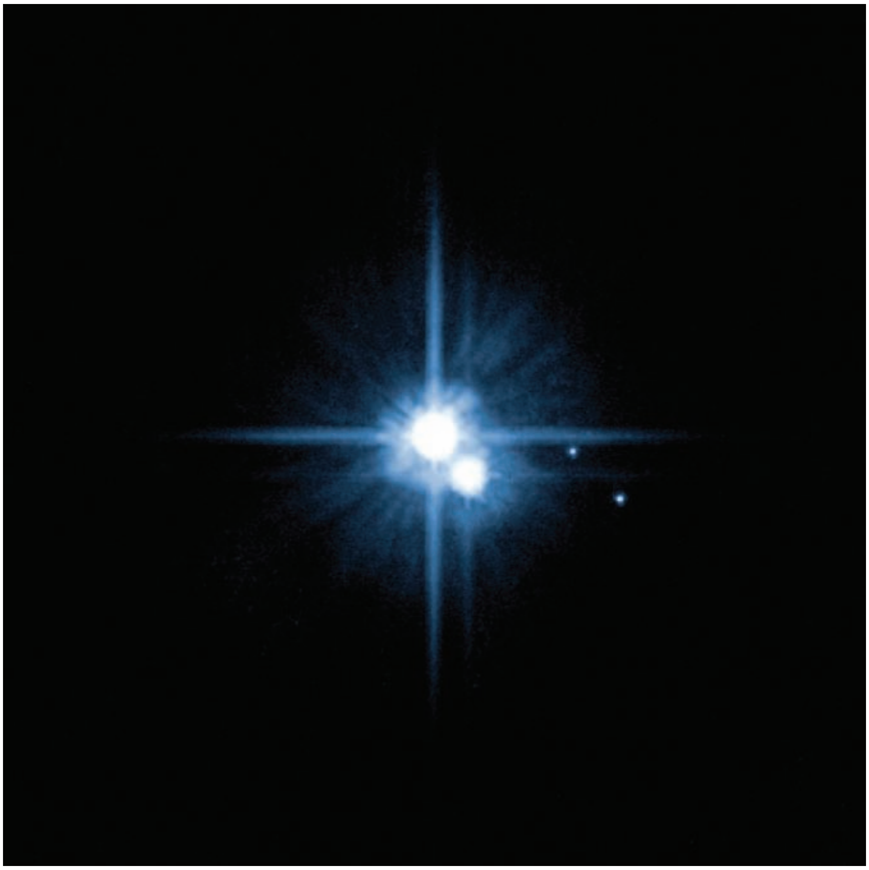


Fig. 15.4 Pluto's three moons Pluto is the brightest object in this image, centered in the diffraction cross. Its large, nearby moon Charon is seen just below and to the right of Pluto. The other two moons, Nix and Hydra, which appear as small points of light to the right of Charon, are about 5000 times fainter than Pluto. The oversized moon Charon was discovered in 1978, while Nix and Hydra were discovered in mid-2005. Compared to Pluto and its large moon Charon, at 2360 and 1210 kilometers in diameter respectively, Nix (*inner moon*) and Hydra (*outer moon*) are estimated to be only 40 and 160 kilometers across. They are about two to three times farther from Pluto than Charon is. (Courtesy of NASA/ESA/Harold Weaver, JHUAPL, Alan Stern, SwRI, and the HST Pluto Companion Search Team.)

two moons are small, each about 80 kilometers across. They move in circular orbits in Pluto's equatorial plane, as Charon does, but at 48 700 and 65 000 kilometers, compared with the 19 640-kilometer separation between the centers of Pluto and Charon.

Perhaps it is no coincidence that the first letters of Nix and Hydra coincide with those of the spacecraft *New Horizons*, which will arrive at Pluto in 2015 and travel on to study other trans-Neptunian objects that most likely resemble Pluto (Focus 15.1).

15.2 Small cold worlds in the outer precincts of the planetary system

As it turned out, Pluto is only one of a host of small icy worlds revolving about the Sun beyond Neptune's orbit (Fig. 15.5). Perhaps a billion of them are comets, too small to be seen with any telescope. But an estimated 70 000 trans-Neptunian objects are over 100 kilometers in diameter and may be potentially observable. Many of them are located close to the orbital plane of the major planets and concentrated between 38 and 48 AU. Others are scattered into more-inclined orbits and extend to greater distances of up to 200 AU.

Focus 15.1 *New Horizons* mission to Pluto and the Kuiper belt

The *New Horizons* spacecraft was launched on 19 January 2006 toward an encounter with Pluto that is expected in July 2015. Since Pluto is about 9.5 billion kilometers away, *New Horizons* must travel at a speed of about 43 000 kilometers per hour to reach Pluto in 9.5 years, which means that the spacecraft will be traveling too fast to be inserted into orbit around Pluto. It will nevertheless move on to explore other parts of the Kuiper belt in which Pluto is located.

The spacecraft's instruments will study the atmosphere, surface and interior of Pluto. They include an ultraviolet spectrometer to measure gas ingredients, an infrared spectrometer to determine surface composition and structure, a radio instrument to study the composition and temperature of the atmosphere, an optical or visual light telescope for high-resolution images of the surface, and energetic particle and dust detectors.

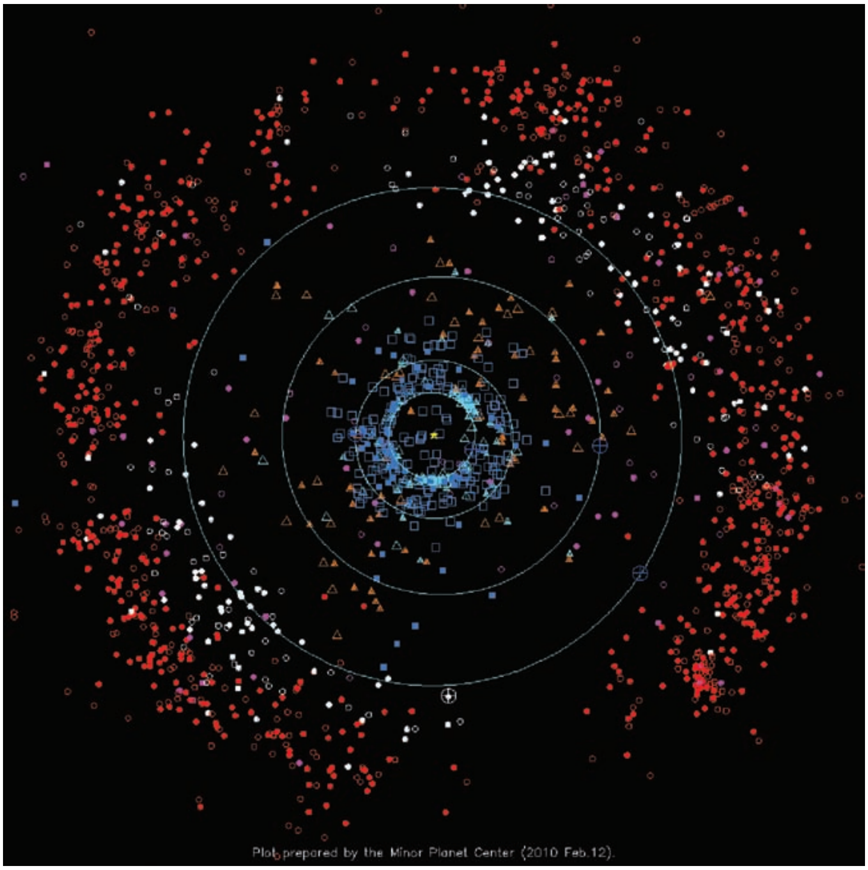


Fig. 15.5 Outer fringes of our planetary realm The orbits of major planets in our solar system are shown in large light blue circles in this plot, prepared on 12 February 2010. The outermost circle denotes the orbit of Neptune and the current location of each planet is marked by a large dark-blue symbol. The current locations of the minor bodies of the outer solar system are shown in different colors to denote different classes of objects. Unusual high-eccentricity objects are shown as cyan triangles, centaur objects as orange triangles, plutoids, a subcategory of dwarf planets, as white circles, and Pluto itself as the large white symbol. Scattered-disk objects are denoted as small magenta circles, and "classical" or "main-belt" objects as small red circles. Objects observed at only one opposition are denoted by open symbols, objects with multiple-opposition orbits are denoted by filled symbols. Filled light-blue squares denote numbered periodic comets; other comets are shown as unfilled light-blue squares. (Courtesy of Gareth Williams, Minor Planet Center.)

The distant, flat ring just outside of Neptune's orbit, at 30 AU, is known as the Kuiper belt in recognition of Gerard P. Kuiper's (1905–1973) prediction of its existence in 1951. He argued that the dark outer edge of the planetary realm is not empty, but is instead full of small, unseen bodies created from the leftover debris of the formation of the giant planets. The low-density material in these distant regions would have been spread out into such a large volume, and moving in such slow, ponderous orbits around the Sun, that it could not gather or coalesce into a body much larger than Pluto.

Jane X. Luu (1963–) and David C. Jewitt (1958–) discovered the first trans-Neptunian, Kuiper-belt object, other than Pluto, on 30 August 1992 using the University of Hawaii's 2.2-meter (87-inch) reflector on Mauna Kea. An electronic detector attached to a large telescope was required to detect the meager amount of sunlight reflected back from such a small object at such great distances. Within a decade of the first discovery, hundreds of these objects were identified (Fig. 15.5). Every one of them is millions of times fainter than can be seen with the unaided eye.

The dynamic orbits of the scattered disk objects occasionally force them into the inner solar system, becoming first centaurs and then short-period comets. The centaurs are small bodies that orbit the Sun between Jupiter and Neptune, and cross the orbit of one or more giant planets (also see Fig. 15.5). The name of the mythological creatures, part human and part horse, was chosen because the centaurs behave as half asteroid and half comet.

The first object to be recognized as a member of this population was discovered by Charles Kowal (1940–) in 1977. It was designated as 2060 Chiron, the two-thousand-and-sixtieth numbered asteroid. Chiron has a diameter of about 200 kilometers, and moves between the orbits of Neptune and Saturn. But then it was discovered that Chiron has a dual personality, developing a luminous shroud of gas and dust like a comet. At least three of the dozens of currently known centaurs display detectable comet comas. Any centaur that is perturbed close enough to the Sun is expected become a comet.

The important thing is that an entirely new class of small worlds has been discovered orbiting the Sun in the outer fringes of the planetary disk, making Pluto much less unique than had been thought previously. Since the pioneering work of Luu and Jewitt, several large trans-Neptunian objects have been found (Figs. 15.6, 15.7). There is Eris, discovered in 2003 by Michael E. Brown (1965–) and colleagues, with a diameter of about

Fig. 15.6 Large trans-Neptunian objects Comparison of the disk sizes of the eight largest trans-Neptunian objects known in 2009. Five of them are known to have moons. The top four were officially designated as dwarf planets, and they have now been re-designated as plutoids. The other four are candidates for this classification. (Courtesy of Wikimedia Commons.)

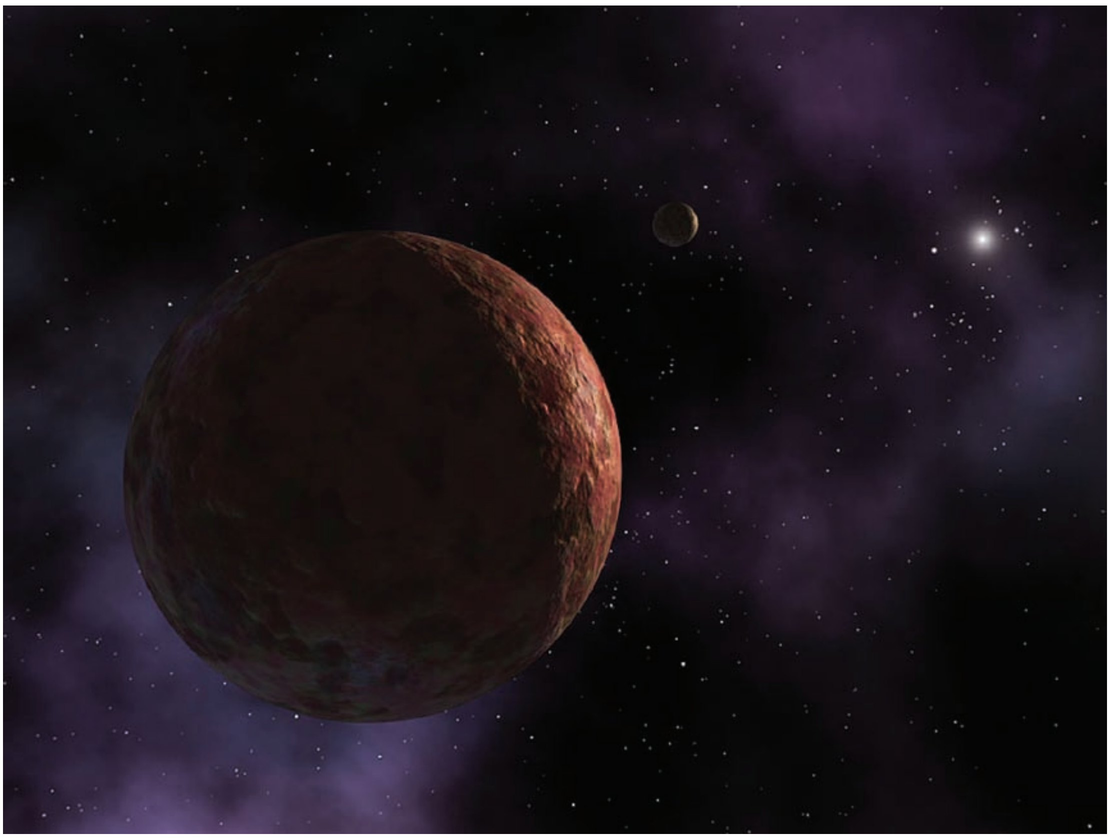


Fig. 15.7 Trans-Neptunian object In this artist's visualization, a small, cold trans-Neptunian object dubbed Sedna is shown where it resides at the outer edges of our solar system, beyond the orbit of Neptune. The object is so far away that the Sun appears as an extremely bright star instead of the large, warm disk observed from Earth. A distant, hypothetical small moon is shown above the object and nearly in the direction of the Sun. (Courtesy of NASA/JPL.)

2400 kilometers. It is larger than Pluto and travels in an eccentric orbit that takes it up to 97 AU from the Sun and well beyond the classical Kuiper belt. Eris is named after the Greek goddess of strife, perhaps as the result of discord over the naming of the object.

Two years later Makemake was found, and named for a Polynesian fertility god. The object is roughly half the diameter of Pluto and orbits the Sun at an average distance of 46 AU. Another large trans-Neptunian object Haumea, named for the Hawaiian goddess of childbirth, has the oblong shape of an egg, and is significantly larger than Pluto in its longest dimension. Haumea has two small moons, and rotates end over end as it moves closer to the Sun than Pluto and then further away.

Since Eris is larger than Pluto, it was initially described as the "tenth planet", but a committee of the International Astronomical Union (IAU) demoted both Pluto and Eris to dwarf planets, different from the eight major ones. They defined a dwarf planet as a celestial body that is in orbit around the Sun and has sufficient self-gravity to overcome rigid body forces and become crushed into a rounded shape. All the major planets meet these criteria, but a dwarf planet has not cleared the neighborhood around its orbit as the major planets have. By this definition, Eris, Haumea, Makemake and Pluto, as well as the largest asteroid, 1 Ceres, are all dwarf planets.

Almost two years after the introduction of this category, the IAU decided to rename all trans-Neptunian dwarf planets as plutoids. So Ceres became the only known dwarf planet that was not also a plutoid, and many of us would prefer to keep on calling it an asteroid.

In any event, the trans-Neptunian objects mark the outer fringes of the planetary system, and this is about where the influence of the Sun's winds stops.

Focus 15.2 The heliosphere's outer boundary

The solar wind carves out a cavity in the interstellar medium known as the heliosphere. The radius of the heliosphere can be estimated by determining the standoff distance, or stagnation point, in which the ram pressure, P_W, of the solar wind falls to a value comparable to the interstellar pressure, P_I. As the wind flows outward, its velocity remains nearly constant, while its density decreases as the inverse square of the distance. The dynamic pressure of the solar wind therefore also falls off as the square of the distance, and we can use the solar-wind properties at the Earth's distance of 1 AU to infer the pressure, P_{WS}, at the stagnation-point distance, R_S. Equating this to the interstellar pressure we have:

$$\begin{aligned} P_{WS} &= P_{1AU} \times \left(\frac{1AU}{R_S}\right)^2 = (N_{1AU}\, V_{1AU}^2) \times \left(\frac{1AU}{R_S}\right)^2 \\ &= P_I \end{aligned}$$

where the number density of the solar wind near the Earth is about $N_{1AU} = 5$ million particles per cubic meter and the velocity there is about $V_{1AU} = 500$ kilometers per second.

To determine the distance to the edge of the solar system, R_S, we also need to know the interstellar pressure, and that is the sum of the thermal pressure, the dynamic pressure, and the magnetic pressure in the local interstellar medium. Its estimated value results in $R_S = 100$ AU or more, well beyond the orbits of the major planets. However, the estimates by different authors give a broad range for the distance to the edge of the heliosphere, depending on the uncertain values of various components of the interstellar pressure, and the termination shock at the edge of the solar wind has now been measured from the *Voyager 1* and *2* spacecraft.

15.3 Edge of the solar system

The wide-open spaces between the planets, once thought to be a tranquil empty void, is swarming with hot, charged invisible pieces of the Sun. They expand and flow away from the Sun at supersonic speeds, faster than 300 kilometers per second, forming a perpetual solar wind. The solar wind is an exceedingly rarefied mixture of electrons and protons set free from the Sun's abundant hydrogen atoms.

The tenuous solar gale moves past the planets and surrounds them, carrying the Sun's magnetic fields and outer atmosphere into the space between the stars. It thereby creates a teardrop-shaped bubble in interstellar space, which is inflated by the solar wind, and threaded by open magnetic fields that have one end attached to the Sun (Fig. 15.8).

How far does the solar wind extend, and where does its influence end? Since the solar wind is weakened by expansion, thinning out as it moves into a greater volume, it eventually becomes too dispersed to repel interstellar forces. The winds are no longer dense or powerful enough to withstand the pressure of gas and magnetic fields coursing between the stars. The radius of this celestial standoff distance, in which the pressure of the solar wind falls to a value comparable to the interstellar pressure, has been estimated at about 100 AU, or one hundred times the mean distance between the Earth and the Sun (Focus 15.2).

Instruments aboard the twin *Voyager 1* and *2* spacecraft, launched in 1977 and now cruising far beyond the

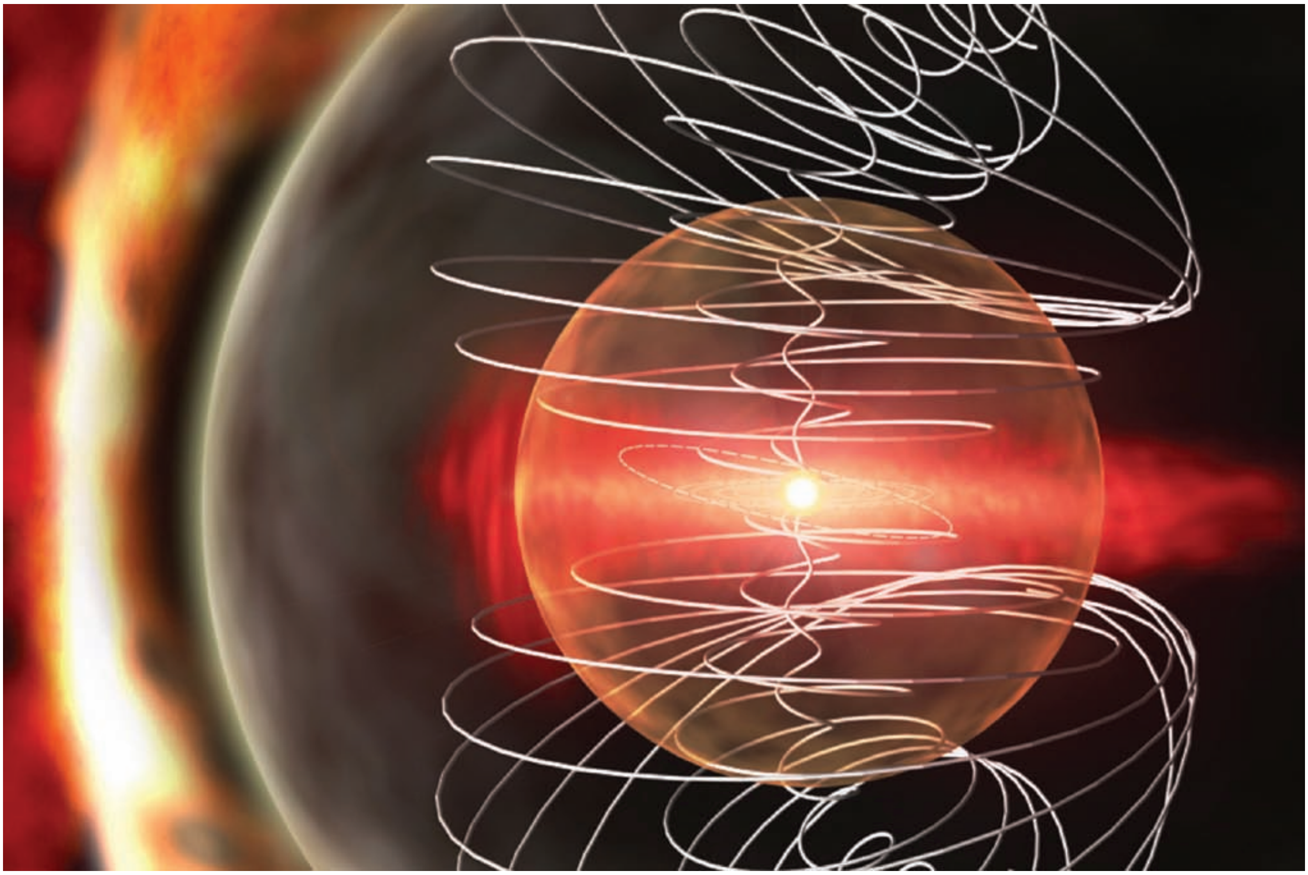


Fig. 15.8 The Sun's domain With its solar wind going out in all directions, the Sun blows a huge bubble within interstellar space called the heliosphere, with the Sun at its center and the planets inside. It is threaded by open magnetic fields that are connected to the Sun at one end, carried into space by the relentless solar wind on the other end, and twisted by the rotating Sun. Interstellar winds mold the heliosphere into a non-spherical teardrop shape, creating a bow shock (*left*) where interstellar forces encounter the solar wind. The heliosphere extends to about 100 times the distance between the Earth and the Sun. (Courtesy of Thomas H. Zurbuchen.)

outermost planets, have approached this edge of the solar system from different directions, *Voyager 1* moving in the northern hemisphere of the heliosphere and *Voyager 2* in the southern hemisphere. *Voyager 1* crossed the termination shock of the supersonic flow of the solar wind on 16 December 2004 at a distance at 94 AU from the Sun. At this distance, the spacecraft's instruments recorded a sudden increase in the strength of the magnetic field carried by the solar wind, as expected when the solar wind slows down and its particles pile up at the termination shock.

Voyager 2 crossed the termination shock on 30 August 2007 at a distance of 84 AU from the Sun. It appears that there is a significant north/south asymmetry in the heliosphere, likely due to the direction of the local interstellar magnetic field.

Both *Voyager 1* and *2* have therefore now crossed into the vast, turbulent heliosheath, the region where the interstellar gas and solar wind interact, due to the reflection and deflection of the solar-wind ions by the magnetized wind beyond the heliosheath. In technical terms, the solar-wind ions in the heliosheath are deflected by magnetosonic waves reflecting off of the heliopause, causing the ions to flow parallel to the termination shock toward the heliotail.

But the Sun's winds are not alone in the dark, cold outer fringes of the solar system. About a million million (10^{12}) unseen comets have been hibernating out there in the Oort comet cloud ever since the solar system formed. Modern calculations suggest that there is an inner Oort cloud with an inner edge at around 3000 AU and a density falling off with greater distance. The outer Oort cloud is continuous with this, but is defined to be those objects at distances greater than 20 000 AU and less than about 200 000 AU. So far, the two *Voyager* spacecraft have traveled nowhere near the more remote comet reservoir. Moreover, the comets are so small and widely spaced that it is exceedingly unlikely that the *Voyager 1* or *2* spacecraft will ever encounter even one of them, just as the *Voyagers* passed through the asteroid belt unaffected by the billions of asteroids there.

Fig. 15.9 Stellar bow shock A crescent-shaped bow shock is formed when the material in the fast wind from the bright, very young star, LL Ori (*center*) collides with the slow-moving gas in its vicinity, coming from the lower right. The stellar wind is a stream of charged particles moving rapidly outward from the star. It is a less energetic version of the solar wind that flows from the Sun. A second, fainter bow shock can be seen around a star near the upper right-hand corner of this image, taken from the *Hubble Space Telescope*. Both stars are located in the Orion Nebula; a bright star-forming region located about 1500 light-years from the Earth. (Courtesy of NASA, the Hubble Heritage Team, STScI, and AURA.)

Both *Voyager* spacecraft are equipped with plutonium power sources expected to last until at least 2020 and perhaps 2025. So they ought to eventually measure the heliopause, at the outer edge of the heliosheath. It is the place where interstellar space begins.

In the meantime, the *Interstellar Boundary EXplorer* (*IBEX*) was launched on 19 October 2009. Instruments on this spacecraft, which operates in Earth orbit, detect neutral, or un-ionized, atoms coming from the termination shock and the boundary between the solar wind and interstellar space.

The motion of the interstellar gas, with its own wind, compresses the heliosphere on one side, producing a teardrop-like, non-spherical shape with an extended tail. A bow shock is formed when the interstellar wind first encounters the heliosphere; just as a bow shock is created when the solar wind strikes the Earth's magnetosphere. The graceful arc of a bow shock, created by an interstellar wind, has been detected around the young star LL Orionis (Fig. 15.9), and this brings us to the captivating recent discoveries of planetary systems around nearby stars other than the Sun.

Part 5 Origin of the solar system and extrasolar planets

16 Brave new worlds

- According to the nebular hypothesis, the Sun and planets formed out of a single collapsing, rotating cloud of interstellar gas and dust called the solar nebula. This hypothesis provides a natural explanation for the highly regular pattern of the planet and satellite orbits.
- Conservation of angular momentum in gravitational collapse suggests that the Sun initially rotated much more rapidly than it does now.
- Spiral nebulae were once thought to be young stars enveloped by nascent planetary systems, but they are now known to be distant galaxies, each containing roughly 100 billion stars.
- The youngest stars in our Milky Way Galaxy are surrounded by dusty planet-forming disks, initially discovered by their infrared radiation, and detected in large numbers and great detail by the *Hubble Space Telescope* and the *Spitzer Space Telescope*.
- Vast interstellar clouds of gas and dust are the incubators of large numbers of newborn stars, many of them embedded in the material from which they arose and surrounded by flattened, rotating planet-forming disks.
- At least two planets with a mass comparable to that of the Earth were discovered orbiting a cold, dark pulsar.
- The first unseen planets circling ordinary Sun-like stars were inferred from the tiny, periodic Doppler wavelength shifts of their parent star's spectral lines, caused by the motion of the orbiting planet. They became known as "hot Jupiters", since they revolve unexpectedly close to their star and have masses comparable to that of Jupiter.
- Hundreds of planets have been found circling nearby stars, including many multi-planet systems. These newfound planets are known as extrasolar planets, or exoplanets for short.
- Most of the exoplanets have been indirectly inferred from the miniscule, periodic velocity changes they create in their parent star, but many have also been inferred from the brief, periodic dimming of starlight when they pass in front of, or transit, their star as viewed from Earth.
- The orbital motions of a few exoplanets have been confirmed by direct observation of the moving planets.

- The *Kepler* mission is searching for Earth-size planets that reside in the habitable zone.
- The *COROT* spacecraft and the world's best ground-based telescopes have detected several super-Earths with a mass between those of the Earth and Jupiter.
- The atmospheres of transiting exoplanets are being investigated using the *Hubble Space Telescope*, the *Spitzer Space Telescope* and ground-based telescopes.
- Water vapor, carbon dioxide and methane have been detected in the atmosphere of one hot Jupiter, named HD 189733b.
- Velocity observations with ground-based telescopes have resulted in the discovery of six planets orbiting a nearby Sun-like star, Gliese 581. One of the exoplanets, designated GL 581g, is located in the potentially habitable zone where liquid water could exist on its surface, and has the right mass, of about three times that of the Earth, to hold an atmosphere.
- There is most likely a very large number of habitable, Earth-size planets in the Milky Way galaxy.

16.1 How the solar system came into being

Any successful theory for the origin of the planets and satellites in our solar system must account for the regular arrangement of their orbits. The planets all move in a narrow band across the sky, the zodiac, which implies that the orbits lie nearly in a plane, and they all orbit the Sun in the same direction that our star rotates. These orbital paths are nearly circular, and the equator of the Sun's rotation nearly coincides with the plane of the planetary orbits. The orbits of most of the satellites imitate the planets in being confined to the planet's equatorial plane and revolving about their planet in the same direction that it orbits the Sun.

This regular orbital arrangement of the planets and satellites is not accidental. Even if a million million million (10^{18}) solar systems were made haphazardly and the planets and moons were thrown into randomly oriented orbits, only one of these solar systems would be expected to look like our own. So it is exceedingly unlikely that the planets became aligned by chance.

Although gravity and motion describe the present behavior of the solar system, they cannot explain the remarkable arrangement. Some additional constraints are required, which describe the state of affairs before the planets were formed and set in motion. These initial conditions are provided by the nebular hypothesis, in which the Sun and planets formed out of a single collapsing, rotating cloud of interstellar gas and dust, called the solar nebula. This hypothesis provides a natural explanation for the highly regular pattern of the planet and satellite orbits.

The German philosopher Immanuel Kant (1724–1804) proposed the nebular hypothesis in his book *Allgemeine Naturgeschichte und Theorie des Himmels*, or *Universal Natural History and the Theory of the Heavens*, published in 1755. He pictured an early Universe filled with thin gas that collected into dense, rotating gaseous clumps. One of these primordial concentrations was the spinning solar nebula. Attracted by its own gravity, the nebula fell in on itself, getting denser and denser, until the middle became so packed, so tight and hot, that the Sun began to shine. Meanwhile, the rotation spun the surrounding material out into a flattened disk revolving about the central Sun. The planets formed from swirling condensations in this circumstellar material, which explains qualitatively why all the planets revolve in the plane that coincides with the equator of the rotating Sun (Fig. 16.1).

The French astronomer and writer Pierre Simon Marquis de Laplace (1749–1827) popularized this nebular hypothesis for the origin of the solar system in 1796, extending it to the formation of rings and moons around the planets. According to Laplace's modification, the shrinking Sun shed a succession of gaseous rings, and each ring condensed into a planet. Then each planet, in turn, became a small rotating nebula in which its own family of rings and satellites was born.

Modern versions of the nebular hypothesis provide additional caveats, but the basic tenants of the original idea are still valid. Billions of years ago, a dense interstellar cloud, the spinning solar nebula, collapsed until the Sun began to shine at its center. The planets formed at the same time, within a flattened rotating disk centered on the contracting proto-Sun.

This is the essence of the original nebular hypothesis, which explained qualitatively the fact that the planets and their moons all revolve in the plane that coincides with the equator of the rotating Sun. The highly regular pattern,

Fig. 16.1 Formation of the solar system An artist's impression of the nebular hypothesis, in which the Sun and planets were formed at the same time during the collapse of a rotating interstellar cloud of gas and dust that is called the solar nebula. The center collapsed to ignite the nuclear fires of the nascent Sun, while the surrounding material was whirled into a spinning disk where the planets coalesced. (Courtesy of Helmut K. Wimmer, Hayden Planetarium, American Museum of Natural History.)

which cannot be accidental, is a natural consequence of the rotation and collapse of a solar nebula composed of gas and dust from which the planets were produced.

There are a few details that need to be explained if the nebular hypothesis is correct, such as the present distribution of mass and angular momentum in the solar system. According to the law of conservation of angular momentum, the rotation of a shrinking object will speed up as the radius decreases, so the young Sun must have been spinning much more rapidly than it does now (Focus 16.1).

Also, if the nebular hypothesis is correct, and the whole solar system originated at the same time, then you might expect the planets to have a similar chemical composition to the Sun. The abundance of the elements in the giant planet Jupiter does indeed mimic that of the Sun, with a predominance of the lightest element hydrogen. Unlike the Sun, the Earth is mainly composed of heavier elements, perhaps because the volatile gases near the young Sun were driven away by its powerful winds. In the inner regions of the solar nebula, the higher temperatures would also vaporize icy material that could not condense, leaving only rocky substances to coalesce and merge together to form the terrestrial planets. The modest masses of the terrestrial planets and their proximity to the Sun did not allow them to capture and retain the abundant lighter gases, hydrogen and helium, directly from the solar nebula.

At larger distances, where the solar nebula was cooler, icy substances could condense and combine with heavier ones to form the large, massive cores of the giant planets. These cores eventually became sufficiently massive to gravitationally accrete, accumulate and capture the surrounding hydrogen and helium. The low temperatures at remote distances from the Sun thus enabled the giant planets to retain the abundant light gases and grow even

Focus 16.1 Conservation of angular momentum in the early solar system

During the gravitational collapse of the rotating interstellar cloud, called the solar nebula, the angular momentum is conserved. For a body of mass M, radius R and rotation period P, this means that

$$\text{Angular momentum} = \frac{2\pi\ M\ R^2}{P} = M\,V\,R = \text{constant}$$

where the rotation velocity V is

$$V = \frac{2\pi\ R}{P}$$

Since the Sun contains 99.87 percent of the mass of the solar system, we can assume that the mass remains constant during the collapse of the solar nebula to form the Sun. We might propose, for argument's sake, that the solar nebula had an initial radius R_{sn}, located at about the current edge of the solar system, where interstellar forces are comparable to those within the solar system. It is located at a radius of about 100 AU, or at about 1.5×10^{13} meters. Assuming that the initial rotation velocity of the solar nebula was comparable to the velocities of interstellar matter, or that $V_{sn} \approx 100$ kilometers per second, the initial rotation period P_{sn} would be $P_{sn} = 2\pi R_{sn}/V_{sn} \approx 1.9 \times 10^9$ seconds. Since angular momentum is conserved during collapse, the rotation period of the solar nebula when it had collapsed to the present size of the Sun, with radius $R_{\odot} = 6.96 \times 10^8$ meters, is estimated to be

$$\text{Rotation period of the young sun} = P_{sn}\left(\frac{R}{R_{sn}}\right)^2 = 4.1 \text{ seconds}$$

And even if the initial solar nebula was rotating at a relatively slow speed of just 1 kilometer per second, the young Sun would be rotating with a period of just 4.1×10^4 seconds or 11.4 hours, comparable to Jupiter's rotation period of 9.9 hours.

However, the Sun now has an equatorial rotation period of 25.7 days, or 2.22×10^6 seconds, many times longer than expected. Some process other than gravitational contraction must have slowed the Sun's spin after its birth. One possibility is the action of magnetic fields that could connect the Sun to the distant, slowly rotating material in the surrounding disk and act like brakes to slow the solar rotation. Another possibility is that frictional forces caused mass to move inwards from the disk to the central Sun while transporting angular momentum outward. Whatever the exact explanation, it seems to apply to other newborn stars, for it is the youngest stars that rotate at the fastest speeds, while older ones spin at a slower rate.

To put the problem in another way, most of the mass of the solar system is in the central Sun, while there is about 10 times as much orbital angular momentum in Jupiter as there is rotational angular momentum in the spinning Sun. So a very small fraction of the mass of the solar system has significant angular momentum, while most of the mass has relatively little angular momentum.

bigger, with large masses and low mass densities. But this scenario does not seem to apply to giant planets recently discovered circling nearby stars in close, hot orbits.

16.2 Newborn stars with planet-forming disks

Twentieth-century astronomers have long been on the lookout for planetary systems around stars other than the Sun, and this has led to some happy surprises. It was once thought, for example that the spiral nebulae (Fig. 16.2) represented an early stage in the evolution of stars, with planets forming around their bright centers. Observations of these objects, with the hope of understanding planetary formation, led to the unanticipated discovery of their enormous velocities and eventually to the discovery of the expanding Universe (Focus 16.2).

It wasn't until 1983 that astronomers used instruments aboard the *InfraRed Astronomical Satellite* (*IRAS*) to unexpectedly obtain the first evidence for planet-forming disks. Using technology pioneered by the military to detect the infrared heat of the enemy, the satellite was designed to detect cosmic infrared radiation, which is mainly inaccessible from the ground owing to absorption in the atmosphere. Because of their low temperature, dust particles emit most of their radiation at infrared wavelengths, while radiating no detectable visible light. It is the other way around for the hot stars, which shine brightly in visible light and emit relatively little infrared.

The *IRAS* instruments found the bright infrared glow of dusty clouds, disks and rings circling bright, massive stars such as the brilliant blue-white Vega, as well as less-massive, solar-type stars. In fact, the youngest nearby stars are usually found embedded in the dense clouds of interstellar gas and dust that spawned them.

Focus 16.2 Spiral nebulae and the discovery of the expanding Universe

In the early 20th century, Vesto Slipher (1875–1969) unexpectedly helped us move beyond the stars into the expanding Universe. Working at the Lowell Observatory in Flagstaff, Arizona, he was measuring the rotations of spiral nebulae, whose bright centers were thought to be newborn stars – the surrounding spiral arms had been interpreted as nascent planetary systems. Using a spectrograph camera with a modest 0.61-meter (24-inch) refractor telescope, he found, in 1917, that the outward velocities of 25 spiral nebulae were well in excess of the velocity of any known cosmic object. Almost all of the spiral nebulae were moving away from the Earth, at astonishingly high velocities, up to 1100 kilometers per second. This suggested to Slipher that the spiral nebulae were stellar systems, or "island universes", seen at great distance rather than nearby stars attended by planet formation.

By 1929 Edwin Hubble (1889–1953) showed that the measured distances, established by him using the superb light-gathering power of the 2.5-meter (100-inch) Hooker telescope on Mount Wilson, were roughly correlated with Slipher's velocities. This relationship is now attributed to the expanding Universe, which no one had anticipated at the time Slipher made his measurements. The spiral nebulae are now known as spiral galaxies, each containing roughly 100 billion stars.

Fig. 16.2 The Andromeda nebula At one time, spiral nebulae like Andromeda were thought to be young stars enveloped by protoplanetary material. We now know that the Andromeda nebula, also known as M31, is the nearest large galaxy to our own, located at a distance of about 2.6 million light-years. Both the Andromeda nebula and our Galaxy are spiral galaxies with similar sizes and total masses, containing roughly 100 billion (10^{11}) visible stars. This photograph was taken in visible light with the 5.0-meter (200-inch) telescope at the Palomar Observatory near Pasadena, California. The two smaller elliptical galaxies are at about the same distance as M31, but with only about one-hundredth of its mass. (Courtesy of the Palomar Observatory.)

The *Spitzer Space Telescope* has recently used its powerful infrared vision to detect hundreds of stars with excess infrared radiation, suggesting that they harbor planet-forming disks. The closest disk system to our own, surrounding the star Epsilon Eridani, contains two infrared-emitting belts, one at approximately the same position as the asteroid belt in our solar system, and the second, denser belt between the first one and a more remote ring of icy comets similar to our own Kuiper belt.

The *Hubble Space Telescope* (*HST*) has discovered flattened disks of dust swirling around at least half the young stars in the Orion nebula, shining in reflected visible light. The high resolution and sensitivity of the *HST* have also been used to obtain detailed images of dusty, planet-forming disks surrounding other Sun-like stars, providing insights to the beginnings of our solar system (Fig. 16.3). The flattened, rotating disks encircling other star suggests that the nebular hypothesis applies to them. This material is expected to coalesce into full-blown planets if it hasn't already done so.

16.3 The plurality of worlds

The ancient Greeks imagined that all matter consists of tiny moving particles, both indivisible and invisible, which they called atoms, and that all material things can be created by the coming together of a sufficient number of atoms. The Greek philosopher Epicurus of Samos (276–194 BC) proposed that the chance conglomerations of innumerable atoms, in an infinite Universe, should result in the formation of a multitude of unseen Earth-like worlds.

The Roman poet Lucretius (99–55 BC) also wrote about the plurality of worlds within a Universe without end, declaring that seeds innumerable in number are rushing on countless courses through an unfathomable Universe, making it highly unlikely that our Earth is the only one to have been created and that all those particles outside are accomplishing nothing.

Astronomers eventually showed that the planets in our solar system are revolving around the Sun, and that the Sun is itself one of innumerable stars. These discoveries opened up the possibility that there might be planets orbiting other stars, some possibly inhabited. Such a

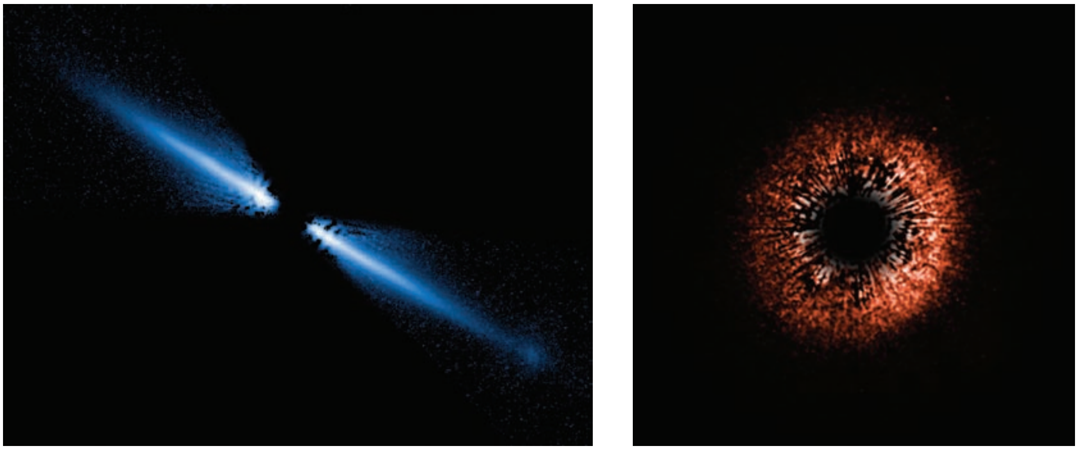


Fig. 16.3 Dusty disks around Sun-like stars Instruments aboard the *Hubble Space Telescope* have obtained these images of the visible starlight reflected from thick disks of dust around two young stars that might still be in the process of forming planets. Viewed nearly face on, the debris disk surrounding the Sun-like star known as HD 107146 (*right*) has an empty center large enough to contain the orbits of the planets in our solar system. Seen edge-on, the dust disk around the reddish dwarf star known as AU Microscopii (*left*) has a similar cleared-out space in the middle. HD 107146 is 88 light-years away, and is thought to be between 50 million and 250 million years old, while AU Microscopii is located 32 light-years away and is estimated to be just 12 million years old. [Courtesy of NASA, ESA, STScI, JPL, David Ardila – JHU (*right*), and John Krist – STScI/JPL (*left*).]

belief dates back at least as far as the Italian philosopher and priest Giordano Bruno (1548–1600), who reasoned in 1584 that these other planets should be orbiting bright stars, and would remain invisible to us because they are much smaller and non-luminous.

Bruno spent the last eight years of his life in the prisons of the Inquisition. He was eventually tried by the Catholic Church, bound to a stake, and burned alive in Rome in 1600, perhaps more for his heretical religious views, such as his doubts about the Immaculate Conception, the Holy Trinity, and Christ's divinity, than for his belief in an infinite Universe filled with countless habitable planets circling other stars.

The invention of the telescope, and the construction of increasingly large ones during ensuing centuries, have enabled astronomers to detect signs of some of these innumerable worlds that were once only imagined. They have caught the light of vast, interstellar clouds of gas and dust, the future incubators of newborn stars (Fig. 16.4).

Some of the giant clouds, with a mass of about a million Suns, are even now in the process of creating stars, falling in on themselves due to the mutual gravitation of their parts. Once this gravitational collapse is underway, the giant cloud fragments into smaller components that eventually collapse to become stars like the Sun.

As the force of gravity pulls the local concentrations of material together, the core of each cloud fragment becomes more compressed and the temperature rises. After about a million years, the central core becomes hot enough to ignite nuclear fusion, and make a star. The most massive stars turn on in a shorter time and shine with greater brightness, lighting up the surrounding material and forming colorful regions within the dark clouds.

A rotating disk is created around a newborn star, the future home of planetary systems. Observations with the *Hubble Space Telescope* and the *Spitzer Space Telescope* indicate that the youngest nearby stars are often embedded in the gas and dust that spawned them, encircled by flattened, rotating planet-forming disks. Instruments aboard the *HST* have discovered flattened disks swirling around at least half the young stars in the Orion Nebula (Fig. 16.5).

But not every interstellar cloud is now in the process of stellar formation. For the most part, the gas is too tenuous and agitated to collapse under its own weight. These dark, stable interstellar clouds contain roughly as much material as is found in visible stars, so there is still plenty of material around to create new stars for billions of years in the future, whenever outside forces might trigger their collapse.

Individual planets shine by reflecting light that is much fainter than the light of the star that illuminates them. The visible light reflected by Jupiter is, for example, about a billion times dimmer than the light emitted by the Sun, and that reflected by Earth is 10 billion times fainter still. As a result, distant planets are almost always too small and too faint to be seen directly in the bright glare of their nearby star. Their presence has only recently been inferred from their miniscule gravitational effects on the motions of the

Fig. 16.4 Mountains of creation The infrared heat radiation of hundreds of embryonic stars (*white/yellow*) and windblown, star-forming clouds (*red*), detected from the *Spitzer Space Telescope*. The intense radiation and winds of a nearby massive star, located just above the image frame, probably triggered the star formation and sculpted the cool gas and dust into towering pillars. (Courtesy of NASA, JPL-Caltech, Harvard-Smithsonian CfA, ESA, and STScI.)

star they revolve around, or when they chance to pass in front of a star, momentarily blocking the star's light when viewed from Earth. Such extrasolar planets, which orbit around stars other than the Sun, are called exoplanets for short.

So the hunt is on, and the prize will be an Earth-size planet with an atmosphere, orbiting a Sun-like star at just the right distance to retain liquid water on its surface. This suggests that we might find companionship in the vast and lonely Universe (Fig. 16.6).

16.4 The first discoveries of exoplanets

Pulsar planets

The first planets to be found outside the solar system were unexpectedly discovered in 1992 by two American radio astronomers, Aleksander Wolszczan (1946–) and Dale A. Frail (1961–). Wolszczan wasn't looking for planets. He was using the giant radio telescope at Arecibo, Puerto Rico, to search for pulsars that spin very rapidly, at the rate of several hundred times a second. Since the telescope is always in high demand, the search only became possible when it was shut down for repairs and pointing in just one direction in the sky. As luck would have it, a previously unknown, fast-spinning pulsar, designated PSR 1257+12 for its position in the sky, happened to pass through the immobile antenna beam. As with other pulsars, it is a tiny, superdense neutron star that emits precisely periodic radio radiation as it rotates.

Once the telescope was repaired, Wolszczan used large computers and comparisons to atomic clocks to measure the arrival of millions and then billions of the rapid, uniform pulses, obtaining a very accurate repetition period.

Fig. 16.5 The great nebula in Orion This nebulosity, the brightest in the sky and designated M 42 or NGC 1976, forms the middle star of Orion's sword. It is 1500 light-years away, a relatively short distance compared with the 100 000 light-year width of our Milky Way Galaxy. Gravity has already pulled some of the interstellar gas into dense concentrations, igniting the celestial fires of newborn stars. The massive stars in the center of the nebula have blown out most of the gas and dust in which they formed, providing a clearer view of other young stars, some of them still embedded in protoplanetary disks in which future planetary systems can form. Hot, massive stars also ionize the nearby interstellar gas, causing this debris of other long-dead stars to fluoresce with red light. More than 3000 stars of various sizes appear in this crisp, detailed image that was taken from the *Hubble Space Telescope* in 2005. (Courtesy of NASA/ESA/Massimo Robberto, STScI and the Hubble Space Telescope Orion Treasury Project Team.)

The best results indicated that the pulsar was spinning once every 0.0062 seconds, or 6.2 milliseconds, rotating on it axis 161 times every second. But there seemed to be something wrong with the data, for the repeating pulses did not match a single, well-defined periodicity, and the computer could not predict exactly when each pulse would arrive at Earth.

The position of the pulsar could be slightly in error, causing a mismatch in the computer analysis of the pulsar timing, which corrects for the motion of the Earth towards

Fig. 16.6 Astrologers of life Two silhouetted figures search the heavens in this 1947 painting by Rufino Tamayo (1889–1991). It may represent our modern attempts to understand the Universe and our search for habitable planets within it. A comet and full Moon illuminate the azure blue sky. The geometric diagrams in the foreground could portray stellar or planetary configurations. A red radio tower in the background sends out signals, perhaps to civilizations on other worlds. (Courtesy of Sotheby Parke Bernet Inc., New York, 1985.)

any point in space. So Dale Frail used the Very Large Array in New Mexico to obtain an accurate location. With the new position, the timing data made sense. The pulse arrival times were being affected by at least two unseen planets.

It was a momentous occasion, for the first planets had been found outside our solar system. But the response of the astronomical community was at best lukewarm, for the planets are orbiting the wrong kind of star. Without any thermonuclear fuel to make it shine, the pulsar PSR 1257+12 is a cold, dark star, emitting no light to warm the newfound planets. In addition, the intense, spinning magnetic fields of the pulsar must accelerate and send lethal high-energy particles and radiation to the planets.

So astronomers were disappointed. They had hoped to discover planets around a perfectly ordinary star like the Sun, whose steady light and heat would at least be compatible with the notion of extraterrestrial life.

Discovery of an unseen planet circling an ordinary star

If any planet in our solar system were placed in orbit around any other star, it would vanish from sight, lost in the star's glare. The presence of the unseen planet has to be instead deduced by indirect means, by recording the way its gravity pulls at the star it orbits. The more massive the planet, and the closer it is to the star, the stronger the planet's gravitational pull on the star and the more the planet perturbs it.

Like two linked, rotating dancers, the planet and star tumble through space, pulling each other in circles. They both orbit a common center of mass where their gravitational forces are equal, somewhat like the equilibrium point of a seesaw, where the forces of two people balance. This fulcrum is closest to the heavier person, or to the massive star in the stellar case. So the star moves in a much smaller circle, a miniature version of the planet's larger path.

To detect this tumbling motion, astronomers had to look for the subtle compressing and stretching of starlight as the unseen planet tugged on the star, pulling it first toward the Earth and then away, causing a periodic shift of the stellar radiation to shorter and then longer wavelengths (Fig. 16.7).

The effect is entirely analogous to the well-known Doppler effect, discovered by the Austrian physicist Christian Doppler (1803–1853), who in 1842 described how sound depends on the relative motion of the source and listener. If the source is moving toward us, the motion compresses the sound waves, pushing them at us and shortening their wavelength. Sounds emitted at a given wavelength then arrive at a shorter wavelength than that emitted by a stationary source; more crests strike the ear each second and a higher frequency or pitch is heard. The sound waves of a receding source are pulled away from us and stretched out to longer wavelengths, with a lower pitch than would be emitted by a non-moving source (see also Section 6.4 and Focus 6.1).

Just as a source of sound can vary in pitch or wavelength, depending on its motion, the wavelength of electromagnetic radiation shifts when the emitting source moves with respect to the observer. If the motion is toward the observer, the Doppler shift is to shorter wavelengths, and when the motion is away the wavelength becomes longer. The amount of the wavelength shift, $\Delta\lambda$, at wavelength λ

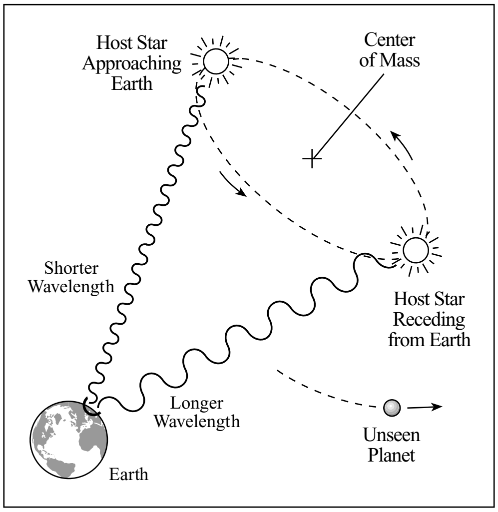


Fig. 16.7 Starlight shift reveals invisible planet An unseen planet exerts a gravitational force on its visible host star. This force tugs the star in a circular or oval path, which mirrors in miniature the planet's orbit. As the star moves through space, it approaches and recedes from Earth, changing the wavelength of the starlight seen from Earth through the Doppler effect. When the planet pulls the star toward us, its light waves pile up in front of it slightly, shortening or "blueshifting" the wavelength we detect. When the planet pulls the star away from us, we detect light waves that are stretched or redshifted. During successive planet orbits, the star's spectral lines are periodically shortened and lengthened, revealing the presence of the planet orbiting the star, even though we cannot see the planet directly.

can be used to infer the velocity of motion along the line of sight, known as the radial velocity V_r, by the equation

$$\frac{\Delta\lambda}{\lambda} = \frac{V_r}{c},$$

where $c = 2.9979 \times 10^8$ meters per second is the velocity of light.

But an orbiting planet produces an exceedingly small variation in the wavelength of spectral lines emitted from its star. Massive Jupiter, for example, makes the Sun wobble at a speed of only about 12 meters per second. To detect the Doppler effect of a star moving with this speed, astronomers would have to measure the wavelengths with an unheard-of accuracy of at least one part in 30 million.

So the effect could only be detected once very sensitive spectrographs were constructed to precisely spread out the light rays. The enhanced light-collecting powers of electronic charge-coupled detectors were also needed to record the dispersed starlight. And since no single line shift is significant enough to be seen, computer software had to be written to add up all the star's spectral lines, which shift together, combining them over and over again at all possible regularities, or orbital periods, with continued comparison to non-moving laboratory spectral lines.

It took decades for astronomers to develop these complex and precise instruments. Then, in the 1990s, the time was ripe, and two Swiss astronomers from the Geneva Observatory in Switzerland, Michel Mayor (1947–) and Didier Queloz (1966–), accomplished the seemingly impossible. In April 1994 they attached a new, exquisitely precise, computerized spectrograph to the 1.93-meter (76-inch) telescope at the Observatoire de Haute-Provence in the south of France, and within a year and a half they had found the first planet that orbits an ordinary star, the faintly visible, Sun-like star 51 Pegasi, only 50.9 light-years away from Earth in the constellation Pegasus, the Winged Horse.

Hints of the planet were found a year before the announcement of its existence, but the Swiss team had to be very careful. There had been many false planetary discoveries before, and they seemed to have found a giant planet with an unexpectedly short orbital period of just 4.23 days. By way of comparison, Jupiter orbits the Sun once every 11.86 years. Observations were stopped in March 1995 because the star moved too close to the Sun to be seen, and renewed in the first week of July when the two astronomers returned with their families. Armed with a precise prediction of what the spectrograph would show if the unseen planet really existed, they pointed the telescope at 51 Pegasi, and saw exactly what they had hoped for. As Mayor described it, the occasion happened like a dream, a spiritual moment.

Mayor and Queloz submitted a discovery paper to *Nature* magazine the following month, and announced it at a professional meeting in Firenze, Italy, on 6 October 1995. They had detected the back-and-forth Doppler shift of the star's light with a regular 4.23-day period, measured by a periodic change of the star's radial velocity of up to 50 meters per second (Fig. 16.8). To produce such a quick and relatively pronounced wobble, the newfound planet had to be large, with a mass comparable to that of Jupiter, which is 318 times heftier than Earth, and moving in a tight, close orbit around 51 Pegasi (Focus 16.3).

Planets that are closer to a star move around it with greater speed and take less time to complete an orbit, all in accordance with Kepler's third law. Thus, the Earth takes a year or 365 days to travel once around the Sun at a mean distance of one astronomical unit (1 AU), while Mercury, the closest planet to the Sun, orbits our star with a period of 88 days at 0.387 AU. A short orbital period of only 4.23 days meant that the newfound planet is located at a distance of just 0.05 AU from its parent star, or about one-eighth the distance between Mercury and the Sun. Thus, a completely unanticipated planet had been found, rivaling

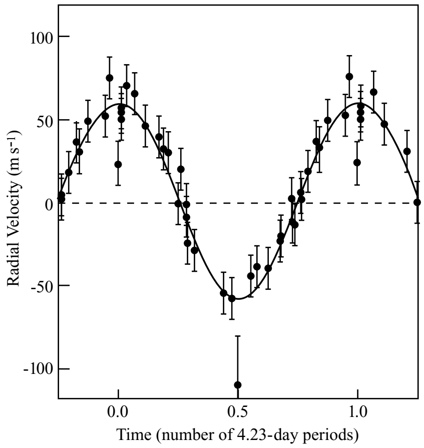


Fig. 16.8 Unseen planet orbits the star 51 Pegasi Discovery data for the first planet found orbiting a normal star other than the Sun. The giant, unseen planet is a revolving around the solar-type star 51 Pegasi, located 50.9 light-years away. The radial velocity of the star, in units of meters per second, has been measured from the Doppler shift of the star's spectral lines. The velocity exhibits a sinusoidal variation with a 4.23-day period, caused by the invisible planetary companion that orbits 51 Pegasi with this period. The observational data (*solid dots*) are fit with the solid line, whose amplitude implies that the mass of the companion is roughly 0.46 times the mass of Jupiter. The 4.23-day period indicates that the unseen planet is orbiting 51 Pegasi at a distance of 0.05 AU, where 1.00 AU is the mean distance between the Earth and the Sun. [Adapted from Michael Mayor and Didier Queloz, "A Jupiter-mass companion to a solar-type star", *Nature* **378**, 355–359 (1995).]

Jupiter in size and revolving around 51 Pegasi in an orbit smaller than Mercury. No one expected a giant planet to be revolving so close to its star.

Less than two weeks after the announcement of a giant planet circling 51 Pegasi, two American astronomers Geoffrey W. Marcy (1955–) and R. Paul Butler (1962–) confirmed the result using the 3-meter (120-inch) telescope at Lick Observatory near Santa Cruz, California. On 17 October 1995, Marcy and Butler issued a press release containing the confirmation, which hit the front-page headlines of newspapers throughout the world. And now that they knew that giant planets could revolve unexpectedly near a star, with short orbital periods, they used powerful computers to re-examine their observations of other nearby stars accumulated during previous years, announcing in January 1996 the discovery of two more Jupiter-sized companions of Sun-like stars.

These were astounding discoveries. In just a few months, astronomers had detected the first planets circling ordinary stars just like our Sun. Other worlds were no longer limited to philosophical musings, scientific speculations, or artists' imaginations. After two millennia, a long dream has come true. We can now look up at the night sky and say that there are definitely invisible planets out there, orbiting perfectly ordinary stars that are now shining brightly in the night sky.

16.5 Hundreds of new worlds circling nearby stars

Scientists had spent decades looking for giant planets far from their central star, only to find that they are easy to detect once you look close in. After scientists realized that a large planet could be so near to its star, they knew where and how to look. And by monitoring thousands of nearby Sun-like stars for years, American and European teams have found more than 400 planets revolving about other nearby stars, most of them massive, Jupiter-size planets.

Some of the newfound worlds travel in circular orbits, like those in the solar system, but much closer to their parent stars than Mercury is to the Sun. Dubbed "hot Jupiters" because of their size and proximity to the intense stellar heat, they are much too hot for life to survive or water to exist. Their temperatures can soar to more than 1000 kelvin, far hotter than the surface of any planet in our solar system. Other newfound planets follow eccentric, oval-shaped orbits that deviate from a circular path, so they venture both near and far from their stars. Many multi-planet systems have also been found as the result of longer and improved observations.

Most of these worlds have been discovered by the wobble they create in the motion of their parent star, but some of them were discovered when they passed in front of the star, causing it to dim, or blink. We haven't mentioned this transit method yet, and it works this way. If the planet happens to have a near edge-on orbit, as seen from Earth, it will periodically cross directly in front of, or transit, its host star. Such a transit can only be seen if the orbit of the distant planet crosses the line of sight from Earth, blocking a tiny fraction of the star's observed light and causing it to periodically dim, over and over again during the planet's endless journey around the star. The size of the planet can be derived from the size of the dip. The planet's temperature can be estimated from the characteristics of the star it orbits and the planet's orbital period.

Some of the transiting hot Jupiters are orbiting in the opposite direction to the rotation of their host star, and others have been found with orbits that are steeply tilted with respect to the star's equatorial plane. In contrast, all eight major planets in our solar system orbit the Sun in the same direction as its rotation and in roughly the same plane, which extends from the Sun's equator. The process that pulled the giant exoplanets so unexpectedly close to

Focus 16.3 Determining the mass and orbital distance of an exoplanet

Planet hunters record the spectral lines of a nearby star, and look for periodic variations in the line-of-sight velocities, V_{obs}, detected from the measured Doppler shifts of the lines. Because the orbital plane is normally inclined to the line of sight, the true orbital velocity, V, is related to the observed velocity by:

$$V_{obs} = V \sin i$$

where i is the inclination angle between the perpendicular to the orbital plane and the line of sight.

The period, P, of the velocity variations is given by Kepler's third law

$$P^2 = \frac{4\pi^2 a^3}{G\,(M_1 + M_2)}$$

where M_1 and M_2 respectively denote the mass of the star and its planet, a is their separation, and the Newtonian gravitational constant G is 6.673×10^{-11} m^3 kg^{-1} s^{-1}. If r_1 and r_2 denote their respective distances from a common center of mass, and we assume circular orbits, then

$$r_1 M_1 = r_2 M_2$$

with

$$a = r_1 + r_2 = \frac{r_1(M_1 + M_2)}{M_2}$$

Since the orbital velocity is

$$V = \frac{2\pi r_1}{P} = \frac{V_{obs}}{\sin i}$$

we can combine this expression with the equation for P and the expression for a to obtain

$$a = \frac{P\,V_{obs}}{2\pi \sin i}\left(\frac{M_1 + M_2}{M_2}\right)$$

Substituting into Kepler's third law, given by the equation for P, gives

$$M_2^3 \sin^3 i = \frac{P\,V_{obs}^3}{2\pi\,G}\,(M_1 + M_2)$$

and since the mass of the star will greatly exceed the mass of the planet, or $M_1 > M_2$,

$$M_2 \sin i \approx \left(\frac{P}{2\pi\,G}\right)^{1/3} V_{obs}\, M_1^{2/3}.$$

For the first exoplanet to be discovered (Fig. 16.8), we have $P = 4.23$ days $= 3.655 \times 10^5$ seconds and $V_{obs} = 50$ meters per second. Under the assumption that $\sin i = 1$ and the star's mass is comparable to the Sun, with $M_1 \approx M_\odot \approx 1.989 \times 10^{30}$ kilograms, we obtain a planet mass of $M_2 = 7.55 \times 10^{26}$ kilograms, which is comparable to the mass of Jupiter, $M_J = 1.9 \times 10^{27}$ kilograms. But the exoplanet is nowhere near as far away from its star as Jupiter is from the Sun, at 7.78×10^{11} m $= 5.2$ AU. The separation, a, is given by

$$a = r_1 + r_2 = r_1\left(1 + \frac{M_1}{M_2}\right) \approx \frac{r_1 M_1}{M_2} = \frac{M_1}{M_2}\frac{P\,V_{obs}}{2\pi \sin i}$$
$$\approx 7.66 \times 10^9 \text{ meters} = 0.05 \text{ AU}$$

which is even closer to the star than Mercury is from the Sun, at 5.79×10^{10} meters or 0.387 AU.

their star may also be responsible for their unanticipated backwards or tilted orbits.

You can keep track of the accelerating pace of discovery at the extrasolar planets encyclopedia at http://exoplanet.eu/ or at http://planetquest.jpl.nasa.gov/. In February 2010, for example, 400 candidate planets and 41 multiple-planet systems had been detected by the radial velocity method and 69 candidate transiting planets had been located.

These have all been indirect detections of exoplanets. The important direct confirmation of a planet circling another star was obtained using the *Hubble Space Telescope* to examine the extensive debris disk of dust surrounding the bright star Fomalhaut; the protoplanetary disk had been discovered in the 1980s using the *InfraRed Astronomy Satellite*. In 2004, an occulting disk was used to block out the star's bright light and enable the space telescope to resolve the visible-light image of a ring of protoplanetary debris, analogous to the Kuiper belt in our solar system. The sharp inner edge of the ring suggested that a nearby planet was clearing out the material beyond it, and by 2008 the light of the Jupiter-size world had been observed. The planet's host star, Fomalhaut, is believed to be a relatively young star; it is only 100 to 300 million years old compared with the Sun's age of 4.6 billion years. Fomalhaut has 2.1 times the Sun's mass and 18 times its luminosity.

Also in 2008, astronomers used the Keck I telescope in Hawaii to directly confirm the orbital motion of three planets around the star HR 8799, using adaptive optics at infrared wavelengths (Fig. 16.9). The host star is roughly 1.5 times as massive as the Sun, and about 5 times as luminous; but it is much younger, with an estimated age of 60 million years. The planets, designated HR 8799 b,

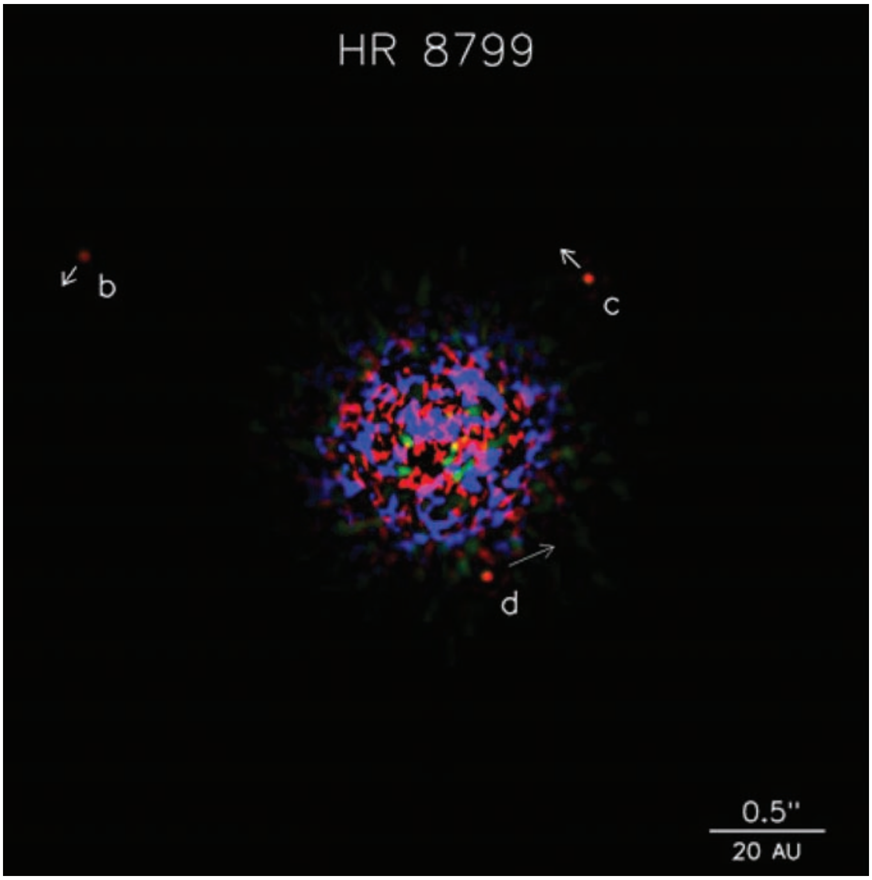


Fig. 16.9 Three planets orbiting a nearby star The speckled residual infrared light of a host star HD 8799 (*center*) and three orbiting planets (*red dots*) are found in this image taken in 2008 from the 10-meter Keck I telescope in Hawaii. The planets all orbit their star in the counter-clockwise direction, as indicated by the arrows, which show their positional displacement between 2004 and 2008. The planets, labeled b, c and d, are located at distances of 70, 40 and 25 AU from the star. The distance of the inner planet is comparable to Neptune's orbital distance at 30 AU. The star HR 8799 has a mass of about 1.5 times that of our own Sun, and is located about 130 light-years away. The planets most likely formed inside a protoplanetary disk about 60 million years ago. The planet masses are estimated as 7, 10 and 10 times that of Jupiter, for b, c and d respectively. (Courtesy of National Research Council, Canada, Christian Marois, Bruce Mcintosh and Keck Observatory.)

c and d, orbit inside a massive dusty disk at distances of roughly twice those of Neptune, Uranus and Saturn from the Sun. Their masses lie between 8 and 10 times the mass of Jupiter. The three planets have subsequently been imaged in infrared light using the relatively small, 1.9-meter diameter portion of the Hale telescope.

Another exoplanet, with a mass of about 9 Jupiter masses, has been observed moving around the star Beta Pictoris, from one side of the star, behind it, and on to the other side (Fig. 16.10). These observations, taken in 2003, 2008 and 2009, were performed using an infrared adaptive optics instrument attached to the Very Large Telescope located in Chile. The relatively cold giant planet, designated Beta Pictoris b, is located at a distance from its host star of between 8 and 15 times the Earth–Sun distance of 1 AU, or at about the same distance as Saturn is from the Sun at 9.539 AU. A dusty disk surrounding Beta Pictoris, extending up to 1000 times the Earth–Sun distance (also see Fig. 16.10), was also discovered during pioneering infrared observations. The presence of a giant planet was subsequently proposed to account for the gap in the dust disk and to explain the observed warp of its inner parts. The planet's host star is just 75 percent more massive than the Sun, but with an estimated age of only 12 million years. Because Beta Pictoris is so young, the exoplanet had to form relatively quickly, in a timespan as short as a few million years.

16.6 Searching for habitable planets

The discovery of hundreds of planets orbiting other stars has created intense excitement and popular interest. From a human perspective, the most interesting planets will be those as small as the Earth, in circular orbits at just the right distance from the heat of a Sun-like star to provide a haven for life.

The orbital size can be calculated from the period of the repeated eclipse, and the planet's temperature estimated. This information would tell us if the planet resides within the warm habitable zone, the range of distances from a star where liquid water can exist on the planet's surface and life might exist. At closer distances, the water would all be boiled away, and at more remote distances it would be frozen solid.

Two missions observe planetary transits from space, attempting to find Earth-size planets in a habitable zone. They are the European *COnvection ROtation and planetary Transits* (*COROT*), mission, launched on 27 December 2006, and NASA's *Kepler* mission, launched on 7 March 2009.

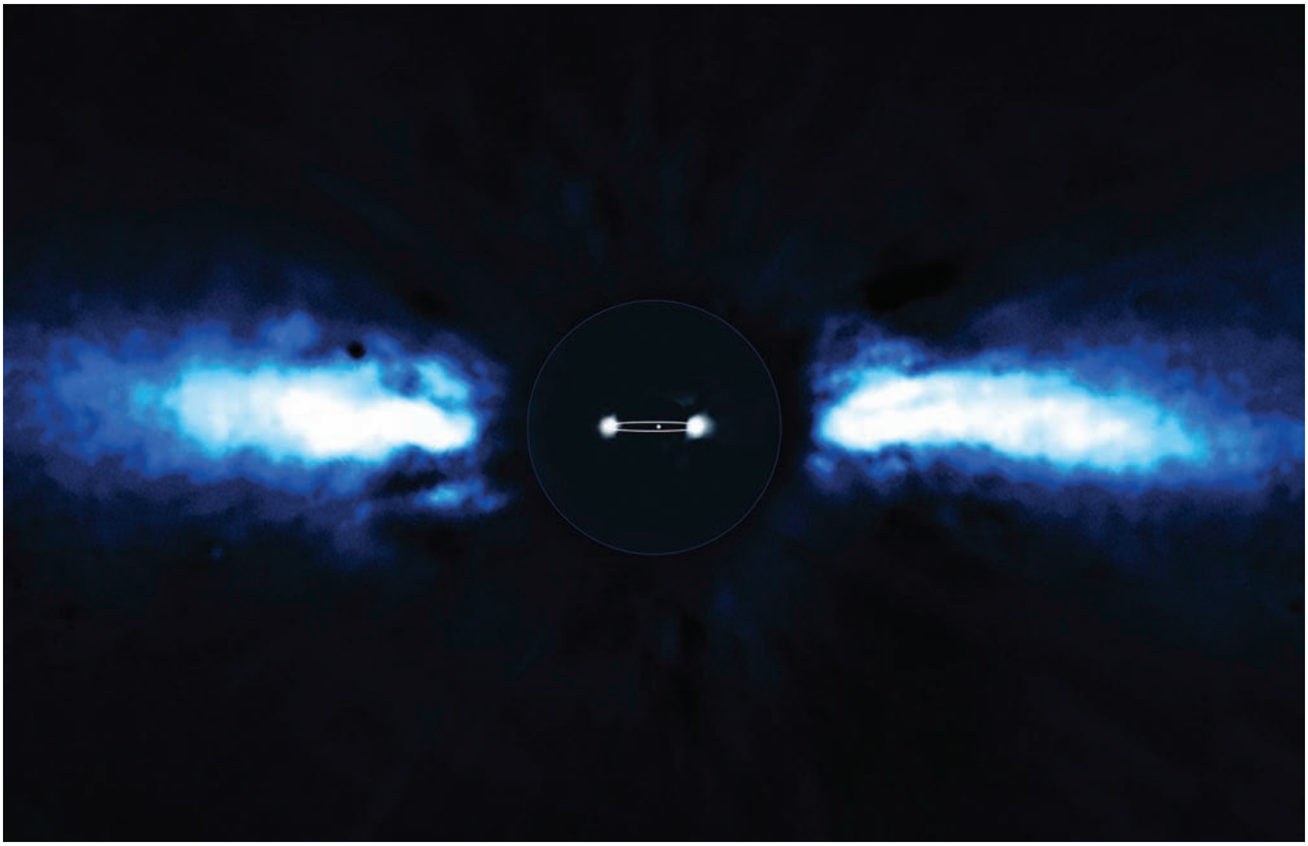


Fig. 16.10 Exoplanet on the move The orbital motion of an exoplanet, denoted by the white elliptical line, was imaged from an adaptive optics instrument attached to the Very Large Telescope in Chile. The small white spot at the center shows the location of the host star, Beta Pictoris. Observations in 2003 are at the left side of the ellipse and those in 2009 are on the right side. The larger dust disc surrounding the host star is also shown by the large flattened blue image at the left and right. (Courtesy of ESO/A.M. Lagrange.)

The *COROT* spacecraft is capable of detecting extrasolar planets with short orbital periods and large terrestrial size. In February 2009, the mission announced its seventh planet discovery named COROT-7b, which has a radius of 1.7 times that of the Earth and is about 4.8 times as massive. The newfound planet's orbital period is nevertheless just 20.5 hours, implying a distance of just 0.017 AU from the host star, or only four times the star's radius and far too hot for comfort.

The *Kepler* mission is specifically designed to detect planets comparable to the Earth in size or smaller, and located at or near the habitable zone. By continuously measuring the brightness of 100 000 stars for four years, it will detect the periodic dimming of starlight produced when the planets pass in front of the star. A transit by an Earth-size planet will produce a small change in the star's brightness of about one ten-thousandth lasting for 2 to 16 hours.

Once detected, the planet's orbital size can be calculated from the period of the repeated eclipse and the mass of the star using Kepler's third law of planetary motion. From the orbital size and the brightness of the star, the planet's temperature can be calculated. This information would tell us if the planet resides in or near the habitable zone. In addition, the size and probable mass of the planet can be found from the depth of the transit, or how much the brightness of the star drops during the transit. The fractional change in brightness, or transit depth, is equal to the ratio of the area of the planet to the area of the star. For the Earth and Sun, as an example, the transit depth is 0.000 084.

Kepler began science operations on 12 May 2009, and within a year had identified more than 700 planet candidates, including at least five candidate systems that appear to exhibit more than one transit. Follow-up observations with ground-based telescopes are required to confirm which candidates are really planets.

Kepler will continue to search for smaller planets at least until November 2012. Since transits of planets in the habitable zone of Sun-like stars occur about once a year and require three transits for verification, it is expected to take at least thee years to locate and verify such a world. Ground-based observations will also be required to confirm the discoveries.

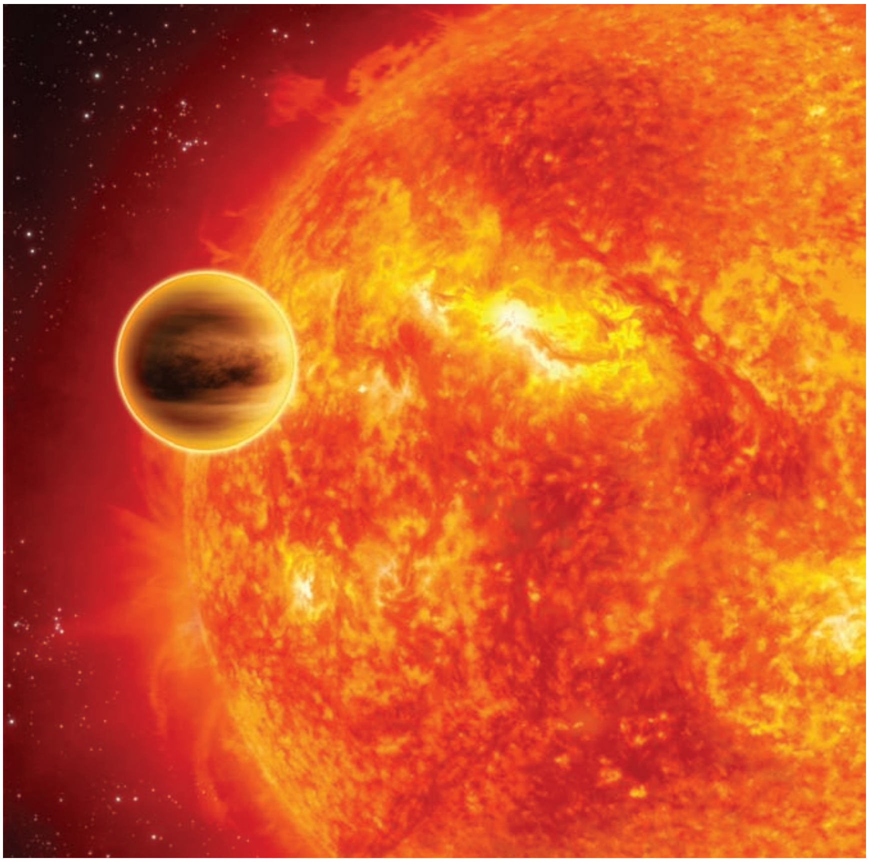


Fig. 16.11 Alien world with an atmosphere An artist's portrayal of a giant Jupiter-size planet passing in front of the star HD 189733b, which lies about 63 light-years away. It is so close to its parent star that it takes just over two days to complete an orbit. In 2007, the spectral signatures of water vapor were observed in the atmosphere of the hot, transiting exoplanet using an instrument aboard the *Spitzer Space Telescope*, and confirmed by observations from the *Hubble Space Telescope*, which also revealed molecules of methane and carbon dioxide. Also in 2007, a spectrograph on the small ground-based InfraRed Telescope Facility, with a mirror diameter of just 3.0 meters, was used to detect carbon doxide and methane in the atmosphere of HD 189733b. Although these organic molecules are also found in living things on Earth, the planet is so massive and so hot that it is considered an unlikely habitat for life. (Courtesy of ESA/C. Carreau.)

In the meantime, the world's best telescopes are being employed to find new exoplanets using the velocity method. The European Southern Observatory's 3.6-meter telescope in La Silla, Chile, has discovered many new ones, including several super-Earths, and the 10-meter Keck I telescope atop Mauna Kea in Hawaii has been used to discover many more, including a super-Earth with about four times the mass of the Earth. Exoplanets with a mass lying between that of the Earth and Jupiter have been dubbed "super-Earths".

The atmospheres of transiting exoplanets are also being investigated using the *Hubble Space Telescope*, the *Spitzer Space Telescope*, and ground-based infrared telescopes. As the planet passes in front and behind its star, astronomers can subtract the light of the star alone, when the planet is blocked, from the light of the star and planet together prior to eclipse. That isolates the emission of the planet and makes possible the detection of the infrared spectral signatures of gases in the planet's atmosphere. Water vapor, carbon dioxide and methane have, for example, been found in the atmosphere of HD 189733b, a hot Jupiter-size planet that orbits its star in just 2.2 days and is nearly 63 light-years away from the Earth (Fig. 16.11).

In 2010 Steven Vogt (1949–) of the University of California, Santa Cruz, and R. Paul Butler (1962–) of the Carnegie Institution of Washington, and their colleagues used 11 years of radial velocity observations of the nearby Sun-like star Gliese 581 (pronounced GLEE-za) to confirm four previously discovered planets orbiting this star, and to discover two more. The sixth planet to be found, designated GL 581g, is within the potentially habitable zone, at the right distance to harbor liquid water on its surface. It has a mass of about three times the Earth's mass, and is thus capable of holding an atmosphere. The dim, red parent star is about 20 light-years away from Earth, about one-third the mass of the Sun, and only about one-hundredth as bright, which means that the habitable zone is relatively close to the star and that a planet in it has relatively short orbital period.

This detection of a potentially habitable planet orbiting a nearby star, coupled with radial velocity surveys of Sun-like stars in the immediate solar neighborhood, suggests that Earth-size exoplanets are common, outnumbering large exoplanets just as sand and small pebbles are more frequent than rocks at the ocean shore. Our Milky Way galaxy is probably teeming with potentially habitable planets located at the right distance to have water and with the right mass to hold an atmosphere. That doesn't mean that they are inhabited with any plant or animal life, but they could be.

Author index

Subject index