Thinking about the Concepts

16. Describe how the giant planets differ from the terrestrial planets.
17. Jupiter's chemical composition is more like that of the Sun than Earth's is. Yet both planets formed from the same protoplanetary disk. Explain why they are different today.
18. What can be learned about a Solar System object when it occults a star?
19. What drives the zonal winds in the atmospheres of the giant planets?
20. Compare the sequence of events in the Process of Science Figure in this chapter with the flowchart of the Process of Science Figure in Chapter 1. Redraw the flowchart, incorporating each of the events leading to the discovery of Uranus as examples in the appropriate boxes.
21. None of the giant planets are truly round. Explain why they have a flattened appearance.
22. What is the source of color in Jupiter's clouds? Uranus and Neptune, when viewed through a telescope, appear distinctly bluish green in color. What are the two reasons for their striking appearance?
23. Which of the giant planets have seasons similar to Earth's, and which one experiences extreme seasons?
24. Jupiter's core is thought to consist of rocky material and ices, all in a liquid state at a temperature of 35,000 K. How can materials such as water be liquid at such high temperatures?
25. Explain how astronomers measure wind speeds in the atmospheres of the giant planets.
26. What is the Great Red Spot?
27. Jupiter, Saturn, and Neptune radiate more energy into space than they receive from the Sun. What is the source of the additional energy?
28. When viewed by radio telescopes, Jupiter is the second-brightest object in the sky. What is the source of its radiation?
29. What creates auroras in the polar regions of Jupiter and Saturn?
30. How might migration of the outer giant planets affect the sizes and orbits of the inner planets?

Applying the Concepts

31. Figure 10.1 shows two different sets of pictures of the outer planets. What is the difference between Figure 10.1a and Figure 10.1b?
32. Figure 10.9d shows the winds on Neptune. The graph, however, does not cover the full planet. Is this likely to mean that the wind speed is zero where there is no white line or that the wind speed is unknown where there is no white line? Explain your reasoning.
33. What creates metallic hydrogen in the interiors of Jupiter and Saturn, and why do we call it metallic?
34. Use Figure 10.15 to estimate the radius of Io's plasma torus in terms of the radius of Jupiter. Convert this value to kilometers, and then look up the answer on the Internet. How close did you get with your simple measurement?
35. The Sun appears 400,000 times brighter than the full Moon in Earth's sky. How far from the Sun (in astronomical units) would you have to go for the Sun to appear only as bright as the full Moon appears in Earth's nighttime sky? How does the distance you would have to travel compare to the semimajor axis of Neptune's orbit?
36. Uranus occults a star at a time when the relative motion between Uranus and Earth is 23.0 km/s. An observer on Earth sees the star disappear for 37 minutes 2 seconds and notes that the center of Uranus passed directly in front of the star.
    a. On the basis of these observations, what value would the observer calculate for the diameter of Uranus?
    b. What could you conclude about the planet's diameter if its center did not pass directly in front of the star?
37. Jupiter's equatorial radius (R_{Jup}) is 71,500 km, and its oblateness is 0.065. What is Jupiter's polar radius (R_{Polar})? (Oblateness is given by $[R_{\text{Jup}} - R_{\text{Polar}}]/R_{\text{Jup}}$.)
38. Ammonium hydrosulfide (NH_4HS) is a molecule in Jupiter's atmosphere responsible for many of its clouds. Using the periodic table in Appendix 3, calculate the molecular weight of an ammonium hydrosulfide molecule, where the atomic weight of a hydrogen atom is 1. (Recall from Working It Out 9.1 that the weight of a molecule is equal to the sum of the weights of its component atoms.)
39. Jupiter is an oblate planet with an average radius of 69,900 km, compared to Earth's average radius of 6,370 km.
    a. Given that volume is proportional to the cube of the radius, how many Earth volumes could fit inside Jupiter?
    b. Jupiter is 318 times as massive as Earth. Show that Jupiter's average density is about one-fourth that of Earth's.
40. The tilt of Uranus is 98°. From one of the planet's poles, how far from the zenith would the Sun appear on summer solstice?
41. A small cloud in Jupiter's equatorial region is observed to be at a longitude of 122.0° west in a coordinate system rotating at the same rate as the deep interior of the planet. (West longitude is measured along a planet's equator toward the west.) Another observation, made exactly 10 Earth hours later, finds the cloud at a longitude of 118.0° west. Jupiter's equatorial radius is 71,500 km. What is the observed equatorial wind speed, in kilometers per hour? Is this wind from the east or west?

42. The equilibrium temperature for Saturn should be 82 K, but the observed temperature is 95 K. How much more energy does Saturn radiate than it absorbs?

43. Neptune radiates 2.6 times as much energy into space as it absorbs from the Sun. Its equilibrium temperature (see Chapter 5) is 47 K. What is its true temperature?

44. Compare the graphs in Figures 10.8a and b. Does atmospheric pressure increase more rapidly with depth on Jupiter or on Saturn? Compare the graphs in Figures 10.8c and d. Does pressure increase more rapidly with depth on Uranus or on Neptune? Of the four giant planets, which has the fastest pressure rise with depth? Which has the slowest?

45. Using Figure 10.8, find the temperature at an altitude of 100 km on each of the four giant planets.

USING THE WEB

46. Go to the *Cassini* website (http://saturn.jpl.nasa.gov). Its final mission is scheduled for late 2016 to 2017. Click on "News." What discovery was reported in a recent news release about Saturn (not about the rings or moons)? Why is this discovery important?

47. Another website for *Cassini* images is found at http://ciclops.org. What do the most recent images of Saturn show? What wavelengths were observed? Are the pictures shown in false color, and if so, why? Why are these images important?

48. a. Go to websites for the NASA *Juno* mission (http://www.nasa.gov/mission_pages/juno and http://missionjuno.swri.edu), a spacecraft that was launched in 2011 and is scheduled to arrive at Jupiter in 2016. What are the science goals of the mission? Examine the mission's trajectory. Why did it loop around the Sun and pass Earth again in 2013 before heading to Jupiter? Why is there a plaque dedicated to Galileo Galilei on the spacecraft?

    b. What are the main instruments for this mission? Are there any data yet? Have any discoveries been reported?

49. Go to the website for the *Voyager 1* and *2* missions (http://voyager.jpl.nasa.gov), which collected data on all four of the giant planets more than two decades ago.

    a. Where are the spacecraft now? Click on "Images & Video." These are still the only close-up images of Uranus and Neptune. What was learned about these planets?

    b. Click on the icon of "The Golden Record," and then on the right, look at scenes, greetings, music, and sounds from Earth. Suppose you were asked to make a new version of the Golden Record, a playlist to send on an upcoming space mission to outside of the Solar System. What would you include in one or more of those categories?

50. Go to the Extrasolar Planets encyclopedia (http://exoplanet.eu/catalog/).

    a. Under "Mass," look for a super-Jupiter planet with a mass significantly larger than that of Jupiter. How far is it from its star—is it a hot Jupiter? Click on the planet name—how was it discovered? If a radius is given, is it more or less dense than Jupiter? Click twice on "Mass" to get a list in descending order—what is the most massive super-Jupiter in the catalog?

    b. Under "Mass," click on "M_{Jup}" so it changes to "M_{Earth}"; do the same under "Radius" so it shows "R_{Earth}." Look for a "Super-Earth." What is its radius? How was it detected? Is there an estimated Mass—if so, what is its density compared with that of Earth? Is it a hot or cold super-Earth?

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

Estimating Rotation Periods of the Giant Planets

digital.wwnorton.com/astro5

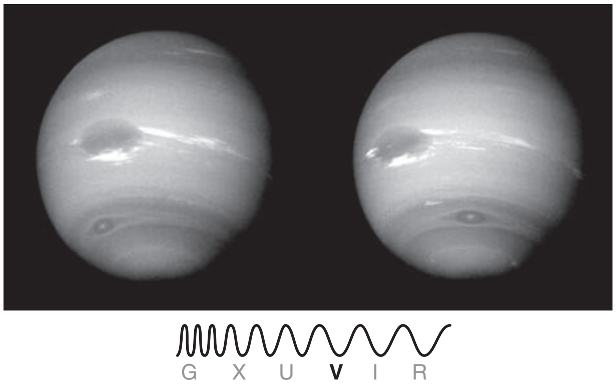


Figure 10.20 These two images of Neptune were taken 17.6 hours apart by *Voyager 2*.

Study the two images of Neptune in **Figure 10.20**. The image on the left was taken first, and the image on the right was taken 17.6 hours later; during this time, the Great Dark Spot completed nearly one full rotation. The small storm at the bottom of the image completed slightly more than one rotation. You would be very surprised to see this result for locations on Earth.

1 What do these observations tell you about the rotation of visible cloud tops of Neptune?

You can find the rotation period of the smaller storm by equating two ratios. First, use a ruler to find the distance (in millimeters) from the left edge of the planet to the small storm in each image (**Figure 10.21a**). The right edge of the planet is not illuminated, so you will have to estimate the radius of the circles traveled by the storms. You can do this by measuring from the edge of the planet to a line through the planet's center (Figure 10.21b). Because the small storm travels along a line of latitude close to a pole, the distance it travels is significantly less than the circumference of the planet.

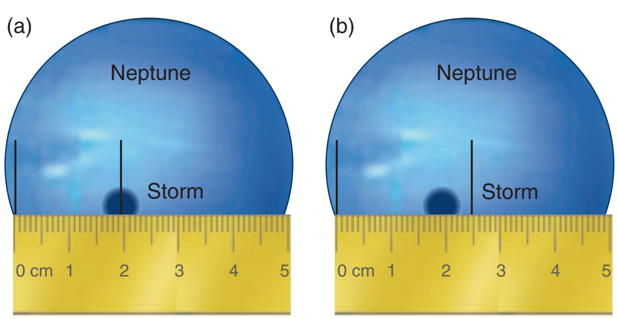


Figure 10.21 (a) How to measure the position of the storm. (b) How to measure the radius of the circle that the storm traveled.

2 Estimate the radius of the circle that this small storm makes around Neptune (in millimeters) by measuring from the edge of the disk to the line through the center of the planet.

---

3 Find the circumference of this circle (in millimeters).

Because the small storm rotated *more than* one time, the total distance it traveled is the circumference of the circle plus the distance between its locations in the two images.

4 Add the numbers you obtained in steps 2 and 3 to get the total distance traveled (in millimeters) between these images.

Now take a ratio and find the rotation period. The ratio of the rotation period, T, to the time elapsed, t, must be equal to the ratio of the circumference of the circle around which it travels, C (in millimeters), to the total distance traveled, D (in millimeters): $T/t = C/D$.

5 You have all the numbers you need to solve for T. What value do you calculate for the small storm's rotation period? (To check your work, note that your answer should be less than 17.6 hours. Why?)

You may be wondering why this calculation works at all. Clearly, the actual distance the small storm traveled is not a small number of millimeters, nor is the circumference of the circle around which it travels. To find the actual distance or circumference, you would multiply both values by the same constant of proportionality. Because you are taking a ratio, however, that constant cancels out, so you might as well leave it out from the beginning.

6 Perform the corresponding measurements and calculations for the Great Dark Spot. What is its rotation period? Think carefully about how to find the total distance traveled, as the Great Dark Spot has rotated *less than* one time around the planet. (To check your work, note that your answer should be more than 17.6 hours. Why?)

---

7 How similar are the rotation periods for these two storms?

---

8 What does this comparison tell you about determining the rotation periods of the giant planets using this method?

---

9 What method do astronomers use instead?

---

11 Planetary Moons and Rings

For centuries, Saturn's rings and the Galilean moons of Jupiter delighted those who looked through telescopes. Since the dawn of the space age, robotic explorers traveling through the Solar System have revealed even more of the diverse collection of moons and rings orbiting other planets.

LEARNING GOALS

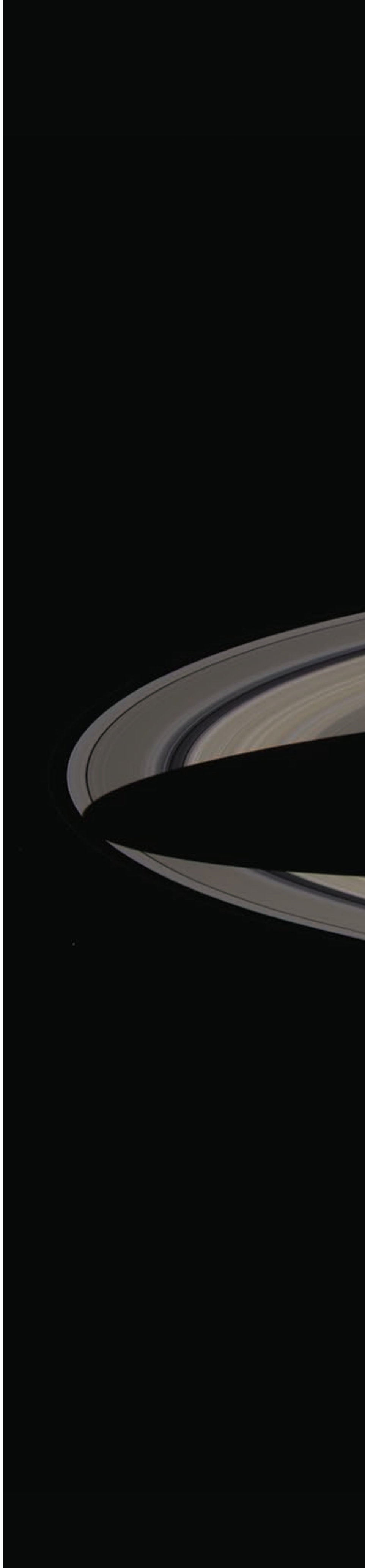
By the conclusion of this chapter, you should be able to:

LG 1 **Compare and contrast the orbits and formation of regular and irregular moons.**

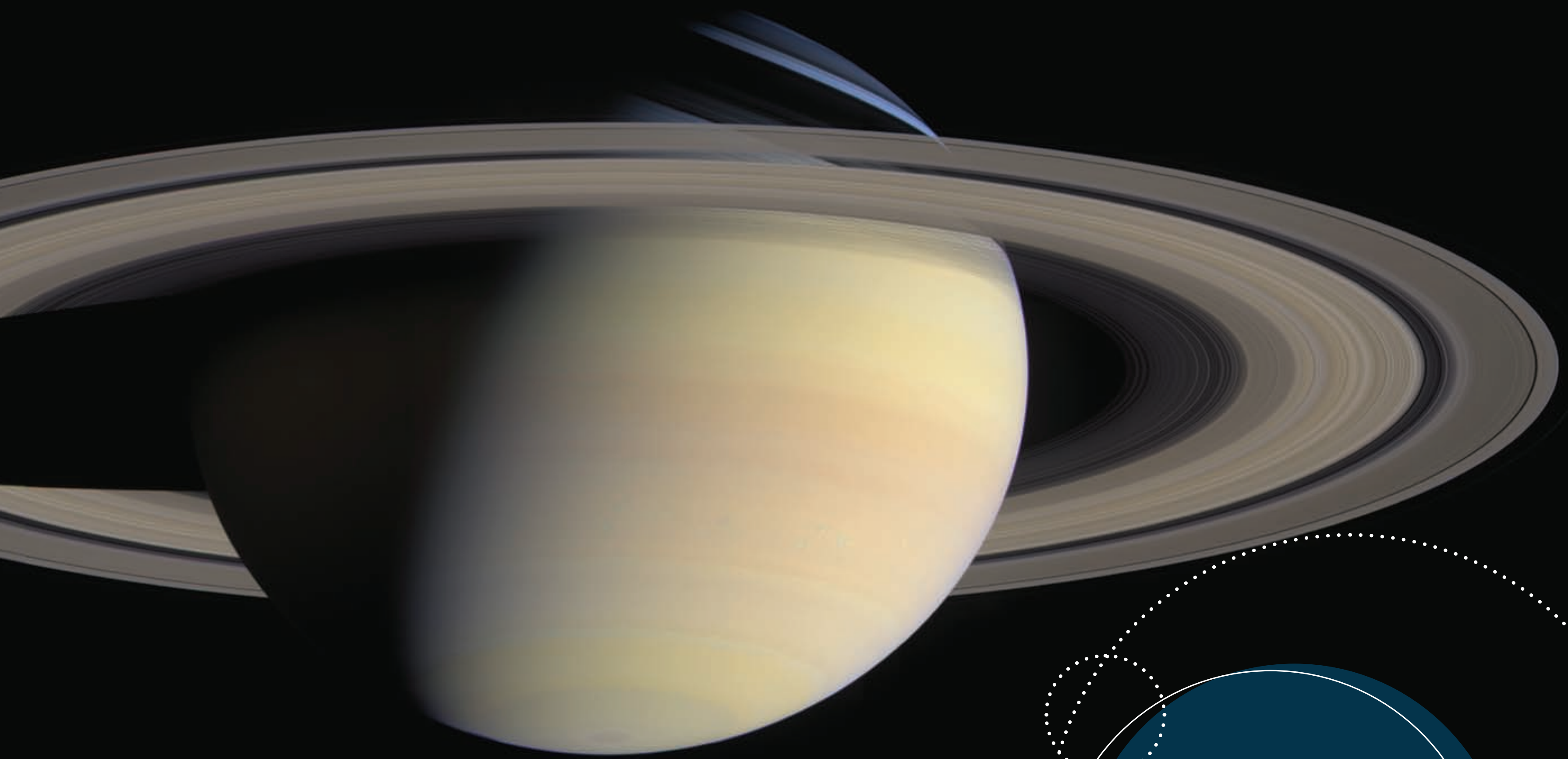
LG 2 **Describe the evidence for geological activity and liquid oceans on some of the moons.**

LG 3 **Describe the composition, origin, and general structure of the rings of the giant planets.**

LG 4 **Explain the role gravity plays in the structure of the rings and the behavior of ring particles.**



Saturn and its rings as viewed by the *Cassini* spacecraft. ▶ ▶ ▶



Why do some planets have rings?

11.1 Many Solar System Planets Have Moons

Most of the planets and some of the dwarf planets in our Solar System have moons. As of 2015, the planets of the Solar System have nearly 150 confirmed moons and a few dozen "provisional" moons that need further confirmation. (Moons around asteroids will be discussed in the next chapter.) New moons in the outer Solar System are still being discovered. Some of these planetary moons are listed in Appendix 4, and an updated list can be found through the "Using the Web" problems at the end of this chapter. Many of these moons are unique worlds of their own, exhibiting geological processes similar to those on the terrestrial planets. Some moons have volcanic activity and atmospheres, and some are likely to contain liquid water under their icy surfaces. Recent discoveries suggest that a few of these moons could have conditions suitable to some forms of life. In this section, we will discuss the orbits and the formation of the moons.

The Distribution of the Moons

The planetary moons of the Solar System are not distributed equally; most are among the giant planets. In the inner part of the Solar System there are only three moons: Earth has one, and Mars has two. Among the dwarf planets, Pluto has five known moons, Haumea has two, and Eris has one. All of the remaining planetary

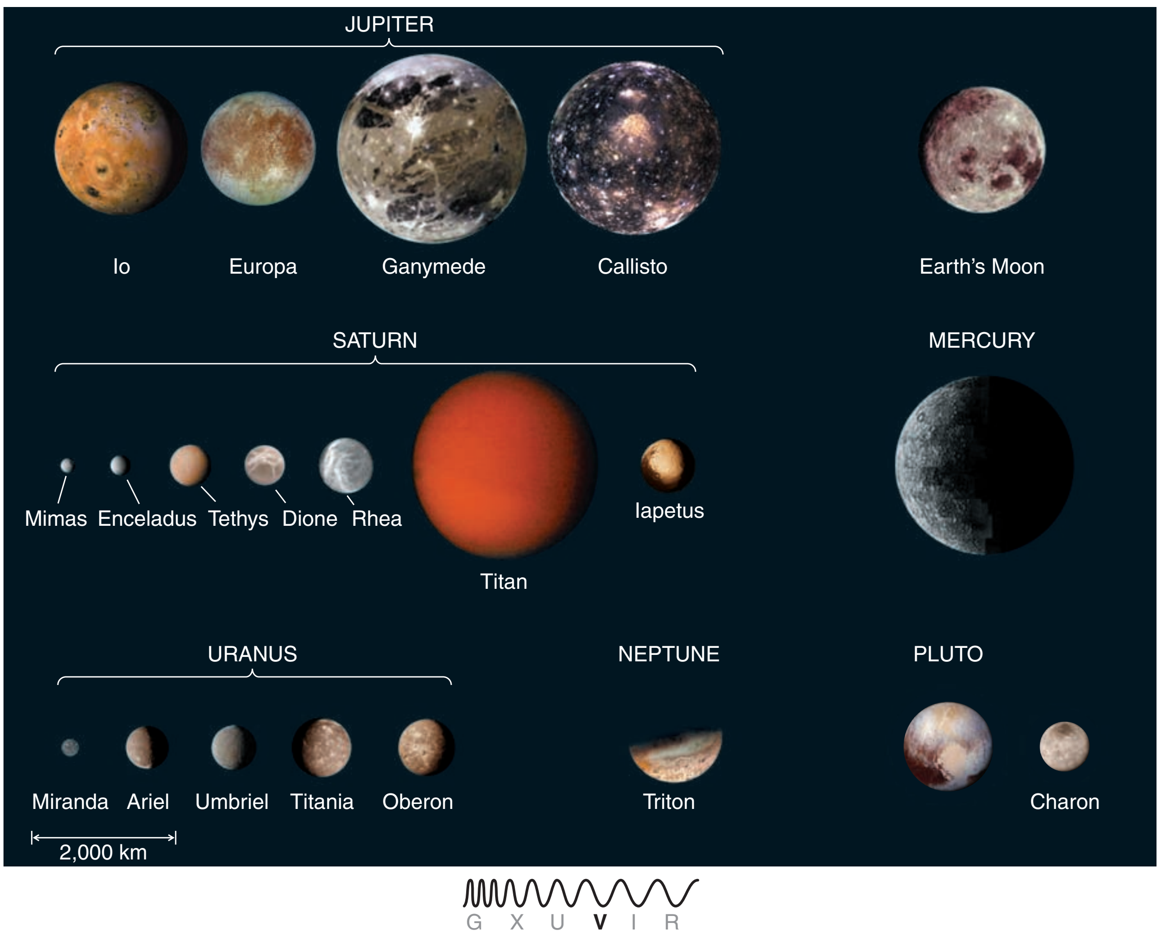


Figure 11.1 This figure shows the major moons of the Solar System, as imaged by various spacecraft. The images are shown to scale. The planet Mercury and dwarf planet Pluto are shown for comparison. The martian moons, Phobos and Deimos, are too small to be shown.

moons belong to the giant planets. Mercury and Venus failed to capture or keep any moons of their own. Earth likely has a moon because of a cataclysmic collision when the planet was young. While the larger planets were forming, they had greater attracting mass and greater amounts of debris around them; consequently, they have more moons.

Figure 11.1 shows the major moons in the Solar System. Some, like Earth's Moon, are made of rock. Others, especially in the outer Solar System, are mixtures of rock and water ice. A few are made almost entirely of ice. Only two moons, Jupiter's Ganymede and Saturn's Titan, are larger in diameter than Mercury, and the smallest known moons are only a kilometer in diameter. Although most moons have no atmosphere, Titan has an atmosphere denser than Earth's, and several have very low-density atmospheres. Scientists suspect that moons accreted from smaller bodies in much the same way that planets accreted from planetesimals, although some may be the product of collisions.

The Orbits of the Moons

Moons can be classified according to their orbits into one of two categories: *regular moons* and *irregular moons*. A **regular moon** lies in its planet's equatorial plane, is close to its planet, and has a nearly circular orbit in the same direction in which its planet rotates (**Figure 11.2a**). About one-third of the moons in the Solar System are regular. These are moons that likely formed from an accretion disk around a host planet at around the time the planet was forming. Our Moon, the Galilean moons of Jupiter, and Saturn's Titan are large, regular moons. With few exceptions, regular moons are tidally locked to their parent planets. Recall from Chapter 4 that tidal locking causes a body to rotate synchronously with respect to its orbit, as Earth's Moon does. When a moon is in synchronous rotation around its planet, the leading hemisphere permanently faces the direction in which the moon is traveling in its orbit around the planet. The trailing hemisphere faces backward. The leading hemisphere is always flying directly into any local debris surrounding the planet, so it may have more impact craters on its surface than the trailing hemisphere.

An **irregular moon** has a more elliptical and more inclined orbit than a regular moon and generally is farther away from its planet than regular moons are from their planets, as shown in Figure 11.2b. Most irregular moons orbit in a direction that is opposite to the rotation of their respective planets; that is, in retrograde orbits. You may recall from Chapter 3 that apparent backward motion of a planet in the sky is called retrograde motion. The largest irregular moons are Neptune's Triton and Saturn's Phoebe. Most of the recently discovered moons of the outer planets are irregular, and many are only a few kilometers across. These are almost certainly bodies that formed elsewhere and were later captured by the planets.

Some of the regular moons have strange orbital characteristics. For example, the moon that is closest to a planet is Phobos, one of the two small moons of Mars (**Figure 11.3**). Phobos is so close that it actually orbits Mars faster than Mars rotates: as seen from Mars, Phobos rises in the west and sets in the east twice a day. It is not known if Phobos and the other moon of Mars, Deimos, were captured from the nearby asteroid belt or if they evolved together with Mars, possibly after a collision early in the history of Mars. Another strange regular moon is Saturn's Hyperion. Hyperion's rotation is chaotic, meaning that it tumbles in its orbit with a rotation period and a spin-axis orientation that are constantly and unpredictably changing. (A *chaotic* system is one in which the final state is exquisitely sensitive to small variations in the initial state. Typically, the result is unpredictable behavior.) No other known moon in the Solar System tumbles like this.

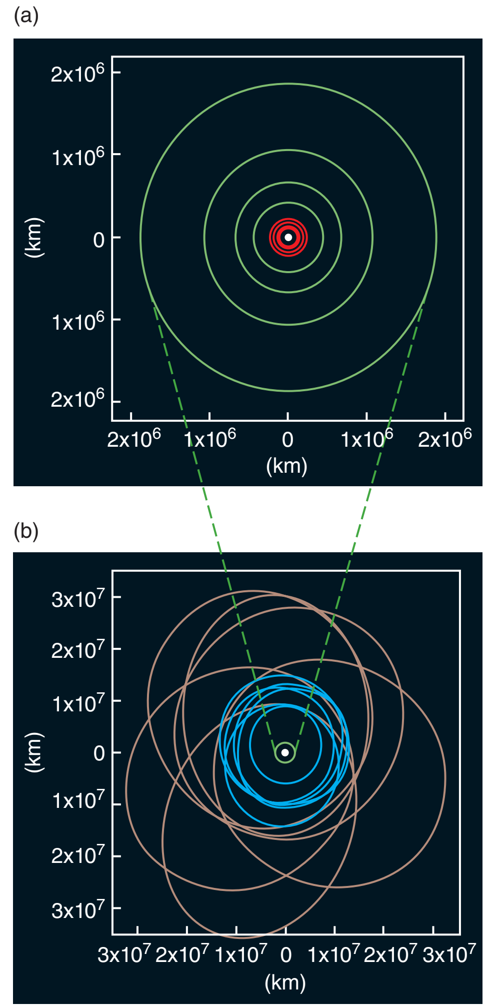


Figure 11.2 These diagrams illustrate the view from "above" the orbits of some of Jupiter's moons. (a) Closer moons, including the Galilean moons, are regular, with nearly circular orbits in Jupiter's equatorial plane. (b) Most of the more distant moons are irregular, with more elliptical, retrograde orbits that are not in the equatorial plane.

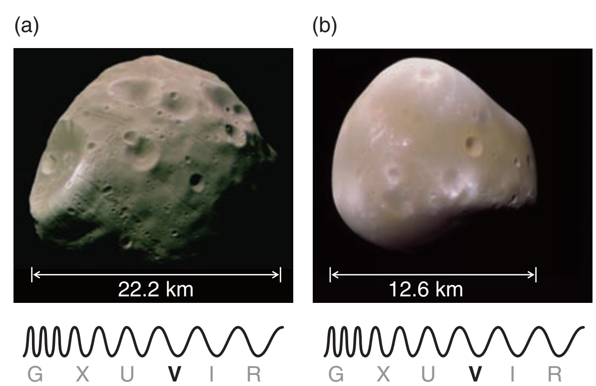


Figure 11.3 These photographs from the *Mars Reconnaissance Orbiter* show the two tiny moons of Mars: (a) Phobos and (b) Deimos.

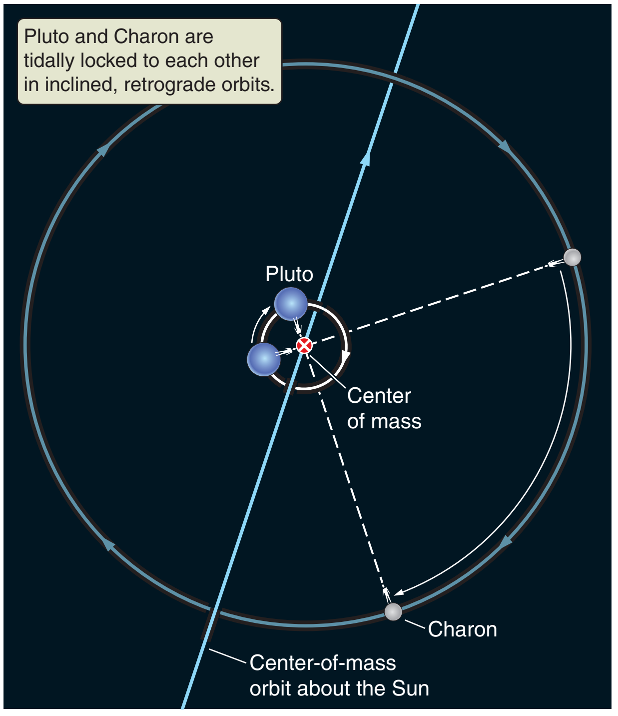


Figure 11.4 This diagram shows the doubly synchronous rotation and revolution in the Pluto-Charon system. The two bodies permanently face one another.

Yet another example of a strange regular moon is Pluto's moon Charon, which is about half as big as Pluto. Pluto and Charon are the only known pair in the Solar System in which both objects are tidally locked to each other. The two are in synchronous rotation, so each has one hemisphere that always faces the other body and another hemisphere that never faces the other body, as shown in **Figure 11.4**. It seems likely that Pluto's highly compact moon system was created by a massive collision between Pluto and another planetesimal, producing a cloud of debris that coalesced to form Charon and the four smaller moons of Pluto, perhaps similar to the way Earth's Moon formed.

Some sets of moons are in synchronized orbits, called **orbital resonances**, where the orbital period of one is a multiple of the orbital period of another. For example, Jupiter's moons Ganymede, Europa, and Io are in a resonance of 1:2:4; that is, for every one orbit of Ganymede, there are two orbits of Europa and four of Io. When the moons are aligned, gravitational effects elongate Io's orbit, which creates variability in the tidal forces of Jupiter on Io. Pluto's five moons are in what appears to be a 1:3:4:5:6 sequence of near resonances. Pairs of some of Saturn's moons are in resonance, and there are also resonances between its moons and gaps in its rings.

The orbits of the moons not only indicate something about their origin. They can also be used to find the masses of their host planets using Kepler's laws, just as you can find the mass of the Sun from the orbital properties of the planets. An example of this is shown in **Working It Out 11.1**.

CHECK YOUR UNDERSTANDING 11.1

Which of the following are characteristics of regular moons? (Choose all that apply.) (a) They revolve around their planets in the same direction as the planets rotate. (b) They have orbits that lie nearly in the equatorial planes of their planets. (c) They are usually tidally locked to their parent planets. (d) They are much smaller than all of the known planets.

11.1 Working It Out Using Moons to Compute the Mass of a Planet

Recall from Working It Out 4.3 that Newton's version of Kepler's law for planets orbiting the Sun could be used to estimate the mass of the Sun. In Chapter 4, we used the following equation to calculate the mass (M) of the Sun:

$$M = \frac{4\pi^2}{G} \times \frac{A^3}{P^2}$$

where A is the semimajor axis of the orbit, and P is the orbital period of any planet.

For moons orbiting a planet, the same equation applies, as long as the moon is much less massive than the planet. Thus, we can use the orbital motion of the moons to estimate the mass of the planet. For example, let's use Jupiter's moon Io, which has an orbital semimajor axis of 422,000 kilometers (km) and an orbital period of $P = 1.77$ days. To match the units in G, we need to put P into seconds:

$$1.77\text{ days} = 1.77\text{ days} \times 24\frac{\text{h}}{\text{day}} \times 60\frac{\text{min}}{\text{h}} \times 60\frac{\text{s}}{\text{min}} = 152{,}928\text{ s}$$

The universal gravitational constant G is equal to $6.67 \times 10^{-20}\text{ km}^3/(\text{kg s}^2)$. Then, the mass of Jupiter is given by

$$M_{\text{Jup}} = \frac{4\pi^2}{G} \times \frac{A^3}{P^2} = \frac{4\pi^2}{6.67 \times 10^{-20}\text{ km}^3/(\text{kg s}^2)} \times \frac{(422{,}000\text{ km})^3}{(152{,}928\text{ s})^2}$$

$$M_{\text{Jup}} = 1.90 \times 10^{27}\text{ kg}$$

You would get the same answer using any other moon of Jupiter.

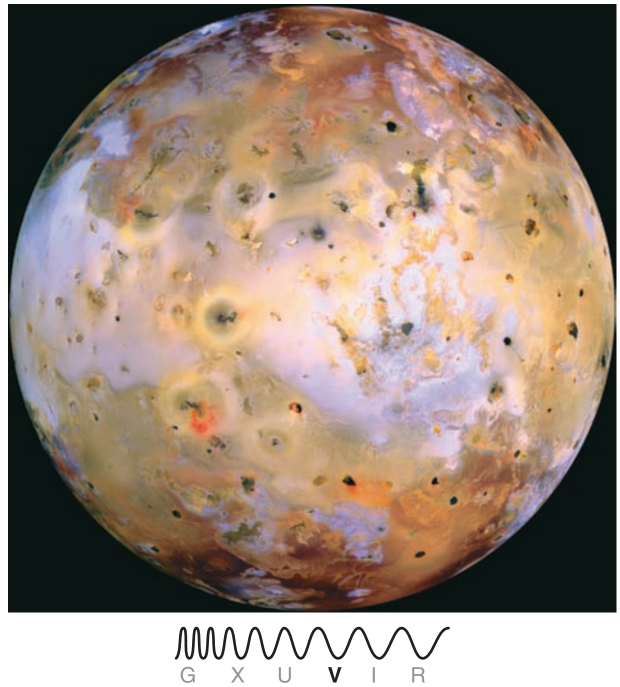
Back before Newton published his law of gravity and before any measured value of the gravitational constant G was possible, Galileo and Kepler showed that P^2/A^3 was the same for each of the four Galilean moons of Jupiter. This demonstrated that Kepler's law applied to systems other than planets orbiting the Sun.

11.2 Some Moons Have Geological Activity and Water

There are several ways to group the moons of the Solar System. Some groupings are based on the sequence of the moons in their orbits around their parent planets; others are based on the sizes or compositions of the moons. In this section, we will organize our discussion of the moons by considering some of the same properties we discussed for the terrestrial planets: the history of the moons' geological activity, and the presence of water and an atmosphere.

Some moons in the Solar System have been frozen in time since their formation during the early history of the Solar System, while others are even more geologically active than Earth. As with the terrestrial planets and Earth's Moon, surface features provide critical clues to a moon's geological history. For example, water ice is a common surface material among the moons of the outer Solar System, and the freshness of that ice indicates the age of those surfaces. Meteorite dust darkens the icy surfaces of moons just as dirt darkens snow late in the season in urban areas. A bright surface often means a fresh surface. The size and number of impact craters indicate the relative timing of events such as volcanism, and this timing enables scientists to gauge whether and when a moon may have been active in the past. Older surfaces have more craters. Observations of erupting volcanoes, which are found on Io and Enceladus, for example, are direct evidence that some moons are geologically active today.

(a)



(b)

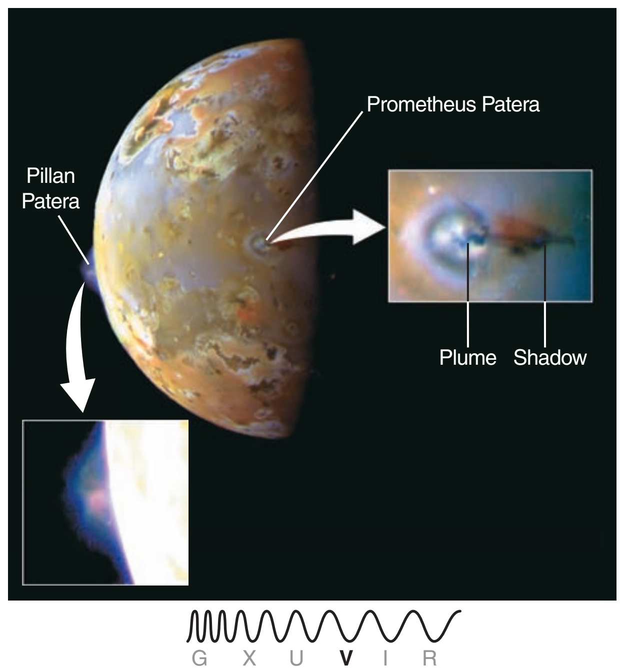


Figure 11.5 (a) This composite image of Jupiter's volcanically active moon Io was constructed from pictures obtained by *Galileo*. (b) The plume from the crater Pillan Patera rises 140 km above the limb of the moon on the left, while the shadow of a 75-km-high plume can be seen to the right of the vent of Prometheus Patera, a volcanic crater near the moon's center.

Io, the Most Geologically Active Moon

One of the more spectacular surprises in Solar System exploration was the discovery of active volcanoes on Io, the innermost of the four large moons of Jupiter. Yet in one of those rare coincidences that happen in science, the changing direction and strength of tidal forces from Jupiter and the nearby moons enabled planetary scientists to predict Io's volcanism just 2 weeks before the moon's volcanic activity was discovered. Why is Io so active? Did you ever take a piece of metal and bend it back and forth, eventually breaking it in half? Touch the crease line, and you can burn your fingers. Just as bending metal in your hands creates heat, the continual flexing of Io generates enough energy to melt parts of its mantle. In this way, Jupiter's gravitational energy is converted into thermal energy, powering the most active volcanism in the Solar System.

Io is just slightly larger than our Moon. Its surface is covered with volcanic features, including vast lava flows, volcanoes, and volcanic craters (**Figure 11.5a**). Lava flows and volcanic ash bury impact craters as quickly as they form, so no impact craters have been observed on the surface. The *Voyager*, *Galileo*, and *New Horizons* spacecraft and the Keck telescope have observed hundreds of volcanic vents and active volcanoes on Io. The most vigorous eruptions spray sulfurous gases and solids hundreds of kilometers above the surface. Some of this material escapes entirely from Io. Ash and other particles rain onto the surface as far as 600 km from the vents. The moon is so active that several huge eruptions often occur at the same time. Figure 11.5b shows the volcanic activity on Io—the source of the material supplying Io's plasma torus and flux tube discussed in Chapter 10.

The surface of Io displays a wide variety of colors—pale shades of red, yellow, orange, and brown. Mixtures of sulfur, sulfur dioxide frost, and sulfurous salts of sodium and potassium likely cause the wide variety of colors on Io's surface. Bright patches may be fields of sulfur dioxide snow. Liquid sulfur dioxide flows

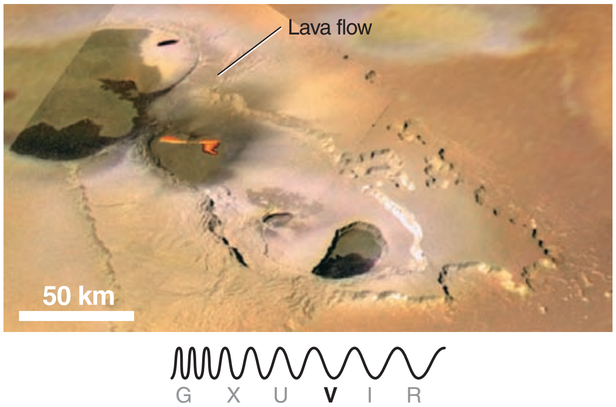


Figure 11.6 This *Galileo* image of Io shows regions where lava has erupted within a caldera. The molten lava flow is shown in false color to make it more visible.

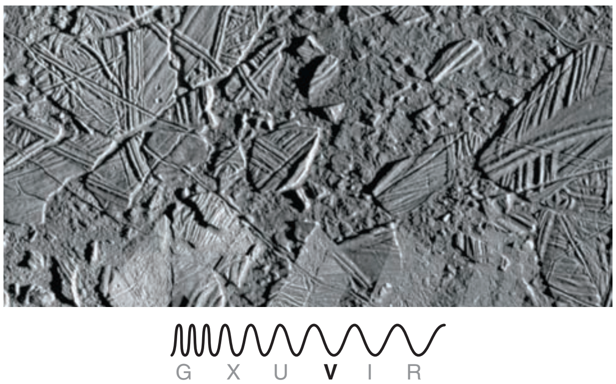


Figure 11.7 This high-resolution *Galileo* image of Jupiter's moon Europa shows where the icy crust has been broken into slabs that, in turn, have been rafted into new positions. These areas of chaotic terrain are characteristic of a thin, brittle crust of ice floating atop a liquid or slushy ocean.

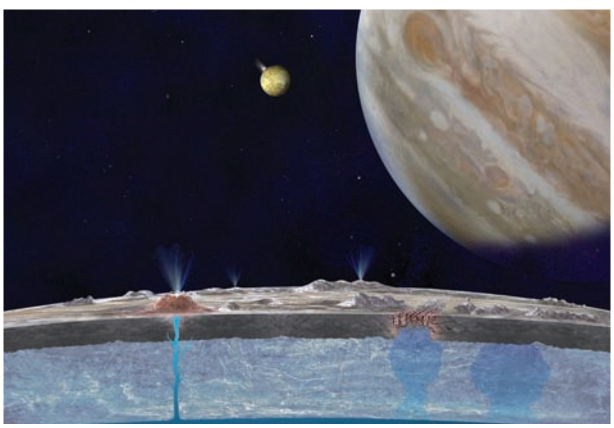


Figure 11.8 This artist's conception of Europa shows liquid bubbling up from the liquid ocean underneath the icy surface. Jupiter and Io are visible in the Europan sky.

beneath Io's surface, held at high pressure by the weight of overlying material. Like water from a spring, this pressurized sulfur dioxide is pushed out though fractures in the crust, producing sprays of sulfur dioxide snow crystals that travel for up to hundreds of kilometers before settling back to the moon's surface. A similar process takes place with a carbon dioxide fire extinguisher. These fire extinguishers contain liquid carbon dioxide at high pressure that immediately turns to "dry ice" snow as it leaves the nozzle.

Spacecraft images reveal the plains, irregular volcanic craters, and flows, all related to the eruption of mostly silicate magmas onto the surface of Io. They also show tall mountains, some nearly twice the height of Mount Everest. Huge structures have multiple summit craters showing a long history of repeated eruptions followed by the collapse of partially emptied magma chambers. Many of the chamber floors are very hot (**Figure 11.6**) and might still contain molten material similar to the magnesium-rich lavas that erupted on Earth more than 1.5 billion years ago. Volcanoes on Io are spread much more randomly than those on Earth, implying a lack of plate tectonics.

Because of its active volcanism, Io's mantle has turned inside-out more than once in the past, leading to chemical differentiation. Volatiles such as water and carbon dioxide probably escaped into space long ago, while most of the heavier materials sank to the interior to form a core. Sulfur and various sulfur compounds, as well as silicate magmas, are constantly being recycled to form the complex surface we see today.

Evidence of Liquid Oceans on Europa and Enceladus

Jupiter's moon Europa is slightly smaller than our Moon and is made of rock and ice and has an iron core. *Voyager* observed an outer shell of water ice with surface cracks and creases. There are few impact craters, so the surface must be young. Regions of chaotic terrain, as shown in **Figure 11.7**, are places where the icy crust has been broken into slabs that have shifted into new positions. In other areas, the crust has split apart, and the gaps have filled in with new dark material rising from the interior. The young surface implies activity, likely powered by continually changing tidal stresses from Jupiter similar to what happens on Io. This activity varies as Europa orbits Jupiter. However, the forces are not as strong on Europa as they are on Io because Europa is farther from the planet (**Working It Out 11.2**).

The *Galileo* spacecraft measured Europa's magnetic field and found that it is variable, indicating an internal electrically conducting fluid. Detailed computer models of the interior of Europa suggest that Europa has a global ocean 100 km deep that contains more water than all of Earth's oceans. This ocean might be salty with dissolved minerals. The lightly cratered and thus geologically young surface indicates that there is energy exchange between the icy crust and liquid water, but it is not known if the icy crust is tens or hundreds of kilometers thick. It is also not yet known if Europa has volcanic activity at the seafloor.

The Hubble Space Telescope may have detected two transient plumes of water vapor erupting from the icy surface, but they were not seen in subsequent observations. The surface is too cold for liquid water, but there might be lakes a few kilometers underneath the surface. Scientists have reanalyzed observations of Europa in light of what has been learned about the numerous subsurface lakes in Antarctica, subglacial volcanoes in Iceland, and ice sheets at both poles of Earth. There may be many shallow "great lakes" underneath the ice, and these would be prime targets for future exploration. **Figure 11.8** is an

11.2 Working it Out Tidal Forces on the Moons

Recall from Chapter 4 that the tidal force between a planet and its moon depends on the masses of the planet and the moon, and the size of the moon, divided by the cube of the distances between them:

$$F_{\text{tidal}} = \frac{2GM_{\text{Jup}}M_{\text{moon}}R_{\text{moon}}}{d^3_{\text{Jup-moon}}}$$

We can use this equation to compare the tidal forces between Jupiter and two of its moons by taking a ratio. For Io and Europa:

$$\frac{F_{\text{tidal-Io}}}{F_{\text{tidal-Europa}}} = \frac{\dfrac{2GM_{\text{Jup}}M_{\text{Io}}R_{\text{Io}}}{d^3_{\text{Io}}}}{\dfrac{2GM_{\text{Jup}}M_{\text{Europa}}R_{\text{Europa}}}{d^3_{\text{Europa}}}} = \frac{\dfrac{M_{\text{Io}}R_{\text{Io}}}{d^3_{\text{Io}}}}{\dfrac{M_{\text{Europa}}R_{\text{Europa}}}{d^3_{\text{Europa}}}}$$

After canceling out the mass of Jupiter M_{Jup} and the constants $2G$, and using data on Io and Europa from Appendix 4—$M_{\text{Io}} = 8.9 \times 10^{22}$ kg, $M_{\text{Europa}} = 4.8 \times 10^{22}$ kg, $R_{\text{Io}} = 1{,}820$ km, $R_{\text{Europa}} = 1{,}560$ km, $d_{\text{Io}} = 422{,}000$ km, and $d_{\text{Europa}} = 671{,}000$ km—we have

$$\frac{F_{\text{tidal-Io}}}{F_{\text{tidal-Europa}}} = \frac{\dfrac{(8.9 \times 10^{22}) \times 1{,}820}{(422{,}000)^3}}{\dfrac{(4.8 \times 10^{22}) \times 1{,}560}{(671{,}000)^3}} = 8.7$$

This comparison shows that the tidal forces on Io are much stronger than those on Europa.

artist's schematic of Europa showing some of the water from the ocean leaking out at the surface, where it would be easier to study. Jupiter and Io are seen in Europa's sky.

Enceladus, one of Saturn's icy moons, shows a wide variety of ridges, faults, and smooth plains. This evidence of tectonic processes is unexpected for a small (500 km) body. The activity on Enceladus is an example of **cryovolcanism**, which is similar to terrestrial volcanism but is driven by subsurface low-temperature liquids such as water and hydrogen rather than molten rock. Some impact craters appear softened, perhaps by the viscous flow of ice, like the flow that occurs in the bottom layers of glaciers on Earth. Parts of the moon have no craters, indicating recent resurfacing. Terrain near the south pole of Enceladus is cracked and twisted (**Figure 11.9a**). The cracks are warmer than their surroundings, suggesting that tidal heating and radioactive decay within the moon's rocky core heat the surrounding ice and drive it to the surface.

Enceladus has a liquid ocean buried beneath 30–40 km of ice crust and is 10 km deep, as an artist has illustrated in Figure 1.10 in Chapter 1. The cracks are warmer than their surroundings, implying that tidal heating and radioactive decay within the moon's rocky core heat ice and drive it to the surface. Active cryovolcanic plumes, like those seen in Figure 11.9b, expel water vapor, tiny ice crystals, and salts. Some of the crystals fall back onto the surface as an extremely fine, powdery snow. The rate of accumulation is very low—a fraction of a millimeter per year—but over time the snow builds up. *Cassini* scientists estimate that the snow may be 100 meters thick in one area near the south pole of Enceladus, indicating that the plume activity has continued on and off for at least tens of millions of years.

Tidal flexing is the likely source of the heat energy coming from Enceladus, as it is on Io and Europa. Enceladus has an orbital resonance with Saturn's moon Dione. A moon made completely of ice would be too stiff for tidal heating to be effective; tidal heating works more effectively with ice over liquid water or cracked ice with some liquid. It remains a mystery why Enceladus is so active while Mimas, a neighboring moon of about the same size—but closer to Saturn and also subject to tidal heating—appears to be geologically dead.

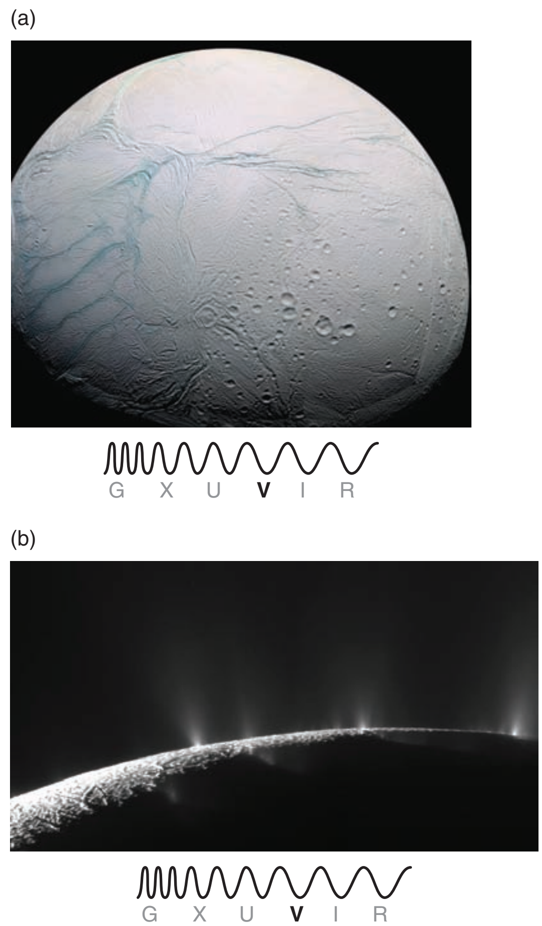


Figure 11.9 These images of Enceladus were taken by *Cassini*. (a) The deformed ice cracks (shown blue in false color) were found to be the sources of cryovolcanism. (b) Cryovolcanic plumes in the south polar region are seen spewing ice particles into space.

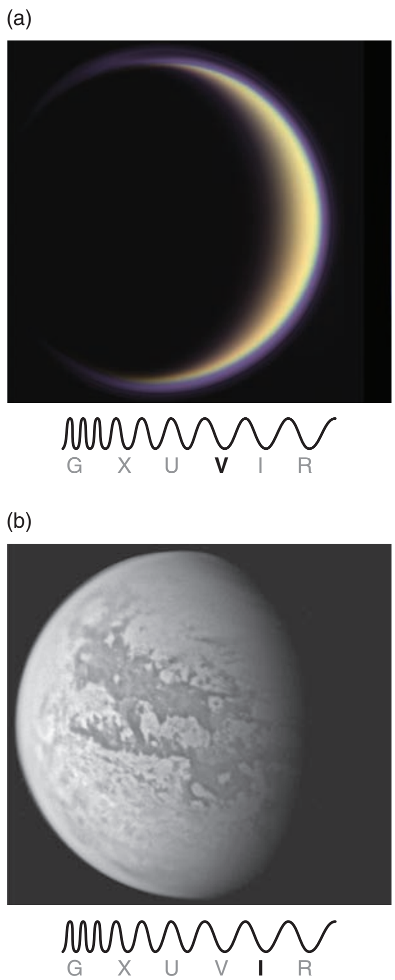


Figure 11.10 These images of Saturn's largest moon, Titan, were taken by *Cassini*. (a) Titan's orange atmosphere is caused by organic, smoglike particles. (b) Infrared-light imaging penetrates Titan's smoggy atmosphere and reveals surface features.

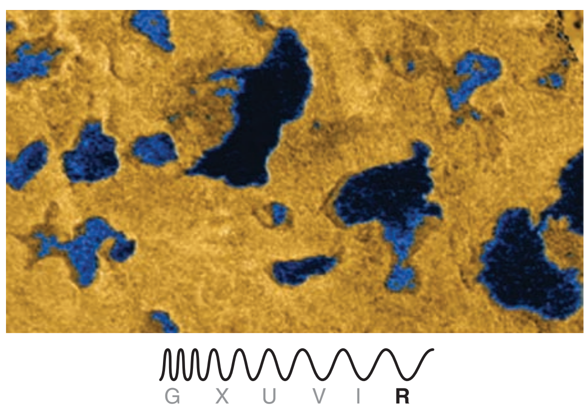


Figure 11.11 Radar imaging (false color) near Titan's north pole shows lakes of liquid hydrocarbons covering 100,000 square kilometers (km^2) of the moon's surface. Features such as islands, bays, and inlets, are visible in many of these radar images.

Titan's Atmosphere and Ocean

Saturn's moon Titan is slightly larger than Mercury and has a composition of about 45 percent water ice and 55 percent rocky material. What makes Titan especially remarkable is its thick atmosphere. Whereas Mercury's secondary atmosphere has been lost to space, Titan's greater mass and distance from the Sun have allowed it to retain an atmosphere that is 30 percent denser than that of Earth. Titan's atmosphere, like Earth's, is mostly nitrogen. As Titan differentiated, various ices, including methane (CH_4) and ammonia (NH_3), emerged from the interior to form an early atmosphere. Ultraviolet photons from the Sun have enough energy to break apart ammonia and methane molecules—a process called **photodissociation**. Photodissociation of ammonia is the likely source of Titan's atmospheric nitrogen. Methane breaks into fragments that recombine to form organic compounds including complex hydrocarbons such as ethane. These compounds tend to cluster in tiny particles, creating organic smog much like the air over Los Angeles on a bad day; this gives Titan's atmosphere its characteristic orange hue (**Figure 11.10a**).

Close-up views of Titan's surface were obtained by the *Cassini* spacecraft. Haze-penetrating infrared imaging showed broad regions of dark and bright terrain (Figure 11.10b). Radar imaging of Titan revealed irregularly shaped features in its northern hemisphere that appear to be widespread lakes and seas of methane, ethane, and other hydrocarbons (**Figure 11.11**). The photodissociative process by sunlight should have destroyed all atmospheric methane within a geologically brief period of about 50 million years, so there must be a process for renewing the methane that is being destroyed by solar radiation. This along with the near absence of impact craters on the surface of Titan suggests recent methane-producing activity. Radar views also indicate an active surface, showing features that resemble terrestrial sand dunes and channels. Heat supplied by radioactive decay could cause cryovolcanism that releases "new" methane from underground. The evidence of active cryovolcanism on Titan is indirect—the presence of abundant atmospheric methane and of methane lakes strongly suggests that Titan has some geological activity.

Titan has terrains reminiscent of those on Earth, with networks of channels, ridges, hills, and flat areas that may be dry lake basins. These terrains suggest a sort of methane cycle (analogous to Earth's water cycle) in which methane rain falls to the surface, washes the ridges free of the dark hydrocarbons, and then collects into drainage systems that empty into low-lying, liquid methane pools. Stubby, dark channels appear to be springs where liquid methane emerges from the subsurface; bright, curving streaks could be water ice that has oozed to the surface to feed glaciers. An infrared camera photographed a reflection of the Sun from such a lake surface. The type of reflection observed proves that the lake contains a liquid and is not frozen or dry. Recent observations might indicate waves on one of these lakes.

Titan is the only moon (aside from Earth's) that has been landed upon. In 2005, *Cassini* released a probe, *Huygens*, which plunged through Titan's atmosphere measuring the moon's composition, temperature, pressure, and wind speeds, and taking pictures as it descended. *Huygens* confirmed the presence of nitrogen-bearing organic compounds in the clouds. During its descent, *Huygens* encountered 120-meter-per-second (m/s) winds and temperatures as low as 88 K. As it reached the surface, though, winds died down to less than 1 m/s and the temperature warmed to 112 K. The pictures taken by *Huygens* showed that the surface was wet with liquid methane, which evaporated as the probe—heated during its passage through the atmosphere—landed in the frigid soil. The surface was also rich with other organic (carbon-bearing) compounds, such as cyanogen and ethane. As shown in **Figure 11.12**, the surface around the landing site is relatively flat and littered with rounded "rocks" of water ice. The dark "soil" is probably a mixture of water and hydrocarbon ices.

As is the case for Saturn's moon Enceladus and Jupiter's moon Europa, gravitational mapping provides indirect evidence that Titan also has an ocean buried beneath its surface. Some of Titan's surface features move by as much as 35 km, which suggests that the crust is sliding on an underlying liquid layer. The current model of Titan is that its rigid ice shell varies in thickness and surrounds an ocean 100 km below the surface, shown in **Figure 11.13**. This ocean would be made of water mixed with dissolved salts—possibly saltier than Earth's Dead Sea. In this model, methane outgassing would occur in hot spots.

Titan is the only moon with a significant atmosphere and the only Solar System body besides Earth that has standing liquid on the surface and a cycle of liquid rain and evaporation. In many ways, Titan resembles a primordial Earth, albeit at much lower temperatures. The presence of liquids and of organic compounds that could be biological precursors for life in the right environment makes Titan another high-priority target for continued exploration.

Figure 11.12 The two water-ice "rocks" just below the center of this Huygens image are about 85 centimeters (cm) from the camera and roughly 15 and 4 cm across, respectively.

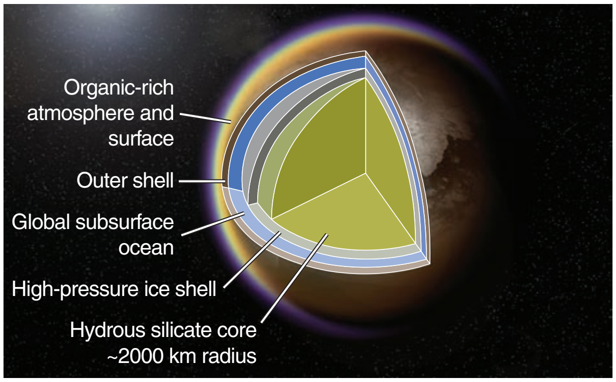


Figure 11.13 This artist's conception of Titan's internal shows how Titan is differentiated, with a core of water-bearing rocks and a subsurface ocean of liquid water. A layer of high-pressure ice surrounds the core, and an outer ice shell is on top of the subsurface ocean.

CHECK YOUR UNDERSTANDING 11.2

Which of the following moons is *not* thought to have an ocean of water beneath its surface? (a) Io; (b) Europa; (c) Enceladus; (d) Titan

Cryovolcanism on Triton

Cryovolcanism also occurs on Triton, Neptune's largest moon. Triton is an irregular moon with a retrograde orbit, which suggests that Triton was captured by Neptune after the planet's formation. As Triton achieved its current circular, synchronous orbit, it experienced extreme tidal stresses from Neptune, generating large amounts of thermal energy. The interior may have melted, allowing Triton to become chemically differentiated.

Triton has a thin atmosphere and a surface composed mostly of ices and frosts of methane and nitrogen at a temperature of about 38 K. The relative absence of craters tells us the surface is geologically young. Part of Triton is covered with terrain that looks like the skin of a cantaloupe (**Figure 11.14**), with irregular pits and hills that may be caused by slushy ice emerging onto the surface from the interior. Veinlike features include grooves and ridges that could result from ice oozing out along fractures. The rest of Triton is covered with smooth volcanic plains. Irregularly shaped depressions as wide as 200 km formed when mixtures of water, methane, and nitrogen ice melted in the interior of Triton and erupted onto the surface, much as rocky magmas erupted onto the lunar surface and filled impact basins on Earth's Moon.

Clear nitrogen ice creates a localized greenhouse effect, in which solar energy trapped beneath the ice raises the temperature at the base of the ice layer. A temperature increase of only 4 K vaporizes the nitrogen ice. As this gas is formed, the expanding vapor exerts very high pressures beneath the ice cap. Eventually, the ice ruptures and vents the gas explosively into the low-density atmosphere. *Voyager 2* found four of these active geyserlike cryovolcanoes on Triton. Each consisted of a plume of gas and dust as much as 1 km wide rising 8 km above the surface, where the plume was caught by upper atmospheric winds and carried for

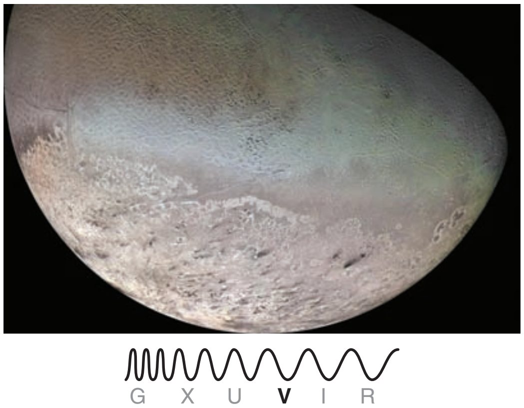


Figure 11.14 This *Voyager 2* mosaic shows various terrains on the Neptune-facing hemisphere of Triton. The lack of impact craters in the "cantaloupe terrain," visible at the top, indicates a geologically younger age than that of the bright, cratered terrain at the bottom.

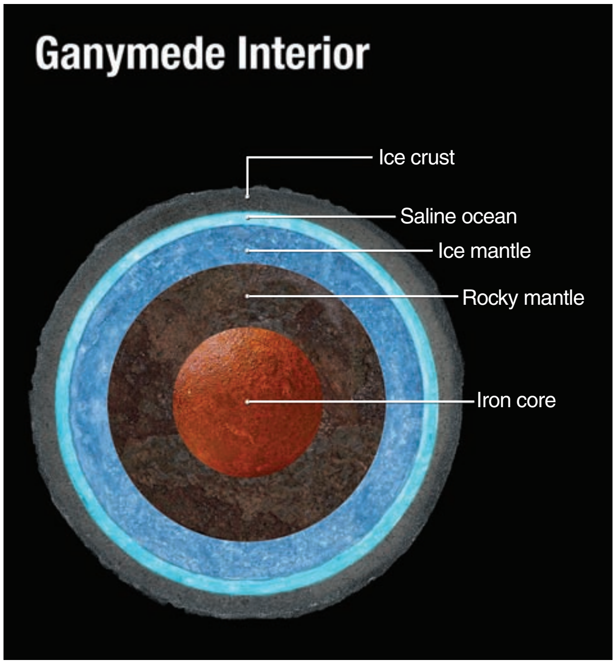


Figure 11.15 This artist's conception of Jupiter's moon Ganymede illustrates that the ocean and ice may be stacked up in multiple layers.

hundreds of kilometers downwind. Dark material, perhaps silicate dust or radiation-darkened methane ice grains, is carried along with the expanding vapor into the atmosphere, from which it subsequently settles to the surface, forming dark patches streaked out by local winds, as seen near the lower right of Figure 11.14.

Formerly Active Moons

Some moons show clear evidence of past ice volcanism and tectonic deformation, but no current geological activity. For example, Jupiter's moon Ganymede, the largest moon in the Solar System, is larger than the planet Mercury. When Jupiter was forming, low temperatures enabled grains of water ice to survive and coalesce along with dust grains into larger bodies at the distance of Ganymede's orbit. In less than half a million years, these bodies accreted to form Ganymede, Jupiter's largest moon. Heating from accretion melted parts of Ganymede so that it is fully differentiated, with outer water layers, an inner silicate zone, and an iron-rich liquid core. As the moon cooled, much of the outer water layer froze, forming a dirty ice crust. Most of the denser materials sank to the central core, leaving an intermediate ice-silicate zone. Ganymede might also have a large, salty ocean underneath its icy surface, maybe 800 km deep, containing 25 times the volume of Earth's oceans (**Figure 11.15**).

Its surface is composed of two prominent terrains: a dark, heavily cratered (and therefore ancient) terrain, and a bright terrain characterized by ridges and grooves. The abundance of impact craters on Ganymede's dark terrain reflects the period of intense bombardment during the early history of the Solar System. The largest region of ancient dark terrain includes a semicircular area more than 3,200 km across on the leading hemisphere. Furrowlike depressions occurring in many dark areas are among Ganymede's oldest surface features. They may represent surface deformation from internal processes or they may be relics of impact-cratering processes.

Impact craters on Ganymede range up to hundreds of kilometers in diameter, and the larger craters are proportionately shallower. The icy crater rims slowly slump, like a lump of soft clay. They are seen as bright, flat, circular patches found principally in the moon's dark terrain (**Figure 11.16**) and are thought to be scars left by early impacts onto a thin, icy crust overlying water or slush (**Figure 11.17**). In Chapter 8, we discussed how planetary surfaces can be fractured by faults or folded by compression resulting from movements initiated in the mantle. On Ganymede, the tectonic processes have been so intense that the fracturing and faulting have completely deformed the icy crust, destroying all signs of older features, such as impact craters, and creating the bright terrain. The energy that powered Ganymede's early activity was liberated during a period of differentiation when the moon was very young. After differentiation was complete, that source of internal energy ran out, and geological activity ceased.

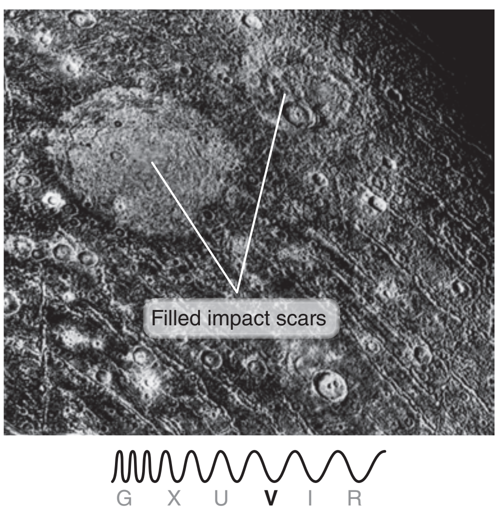


Figure 11.16 This *Voyager* image shows filled impact scars on Jupiter's moon Ganymede.

Many other moons show evidence that they experienced an early period of geological activity that resulted in a dazzling array of terrains. A 400-km impact crater scars Saturn's moon Tethys, covering 40 percent of its diameter, and an enormous canyonland wraps at least three-fourths of the way around the moon's equator. Saturn's moon Dione shows bright ice cliffs up to several hundred meters high, created by tectonic fracturing. The trailing hemisphere of Saturn's Iapetus is bright, reflecting half the light that falls on it, while much of the leading

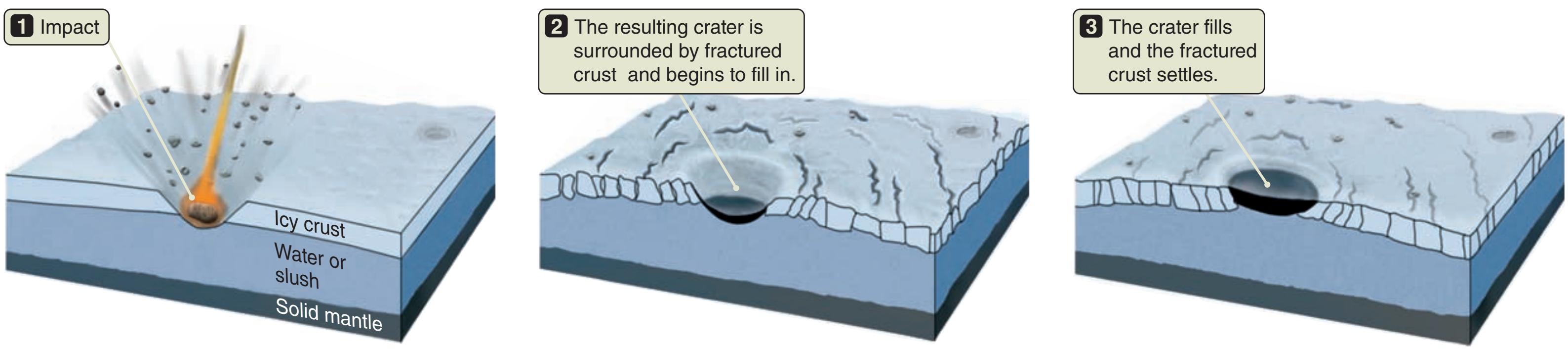


Figure 11.17 Filled impact scars form as viscous flow smooths out structures left by impacts on icy surfaces.

hemisphere is as black as tar. These dark deposits appear *only* in the leading hemisphere of Iapetus, suggesting that they might be debris that was blasted off small retrograde moons of Saturn by micrometeoritic impacts and swept up by Iapetus as it moved along in its prograde orbit around Saturn.

Saturn's moon Mimas, no larger than the state of Ohio, is heavily cratered with deep, bowl-shaped depressions. The most striking feature on Mimas is a huge impact crater in the leading hemisphere. Named "Herschel" after astronomer Sir William Herschel, who discovered many of Saturn's moons, the crater is 130 km across, a third the size of Mimas itself (**Figure 11.18**). It is doubtful that Mimas could have survived the impact of a body much larger than the one that created Herschel. Some astronomers think that Mimas (and perhaps other small, icy moons as well) was hit many times in the past by objects so large as to fragment the moon into many small pieces. Each time this happened, the individual pieces still in Mimas's orbit would coalesce to re-form the moon, perhaps in much the same way that Earth's Moon coalesced from fragments that remained in orbit around Earth after a large planetesimal impacted Earth early in its history.

Areas on Uranus's small moon Miranda have been resurfaced by eruptions of icy slush or glacierlike flows. Other moons of Uranus—Oberon, Titania, and Ariel—are covered with faults and additional signs of early tectonism. On Ariel, in particular, very old, large craters appear to be missing, perhaps obliterated by earlier volcanism.

Geologically Dead Moons

Geologically dead moons, such as Jupiter's Callisto, Saturn's Hyperion, Uranus's Umbriel, and a large assortment of irregular moons, are moons for which there is little or no evidence of internal activity having occurred at any time since their formation. The surfaces of these moons are heavily cratered and show no modification other than the cumulative degradation caused by a long history of impacts.

Callisto is the third largest moon in the Solar System, just slightly smaller than Mercury. It is also the darkest of the Galilean moons of Jupiter, yet it is still twice as reflective as Earth's Moon. This brightness indicates that Callisto is rich in water ice, but with a mixture of dark, rocky materials. Except in areas that experienced large impact events, the surface is essentially uniform, consisting of

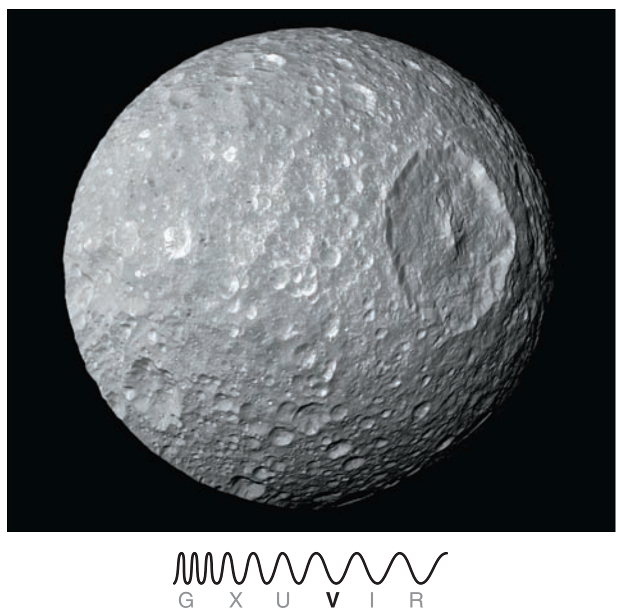


Figure 11.18 This *Cassini* image shows Saturn's moon Mimas and the crater Herschel.

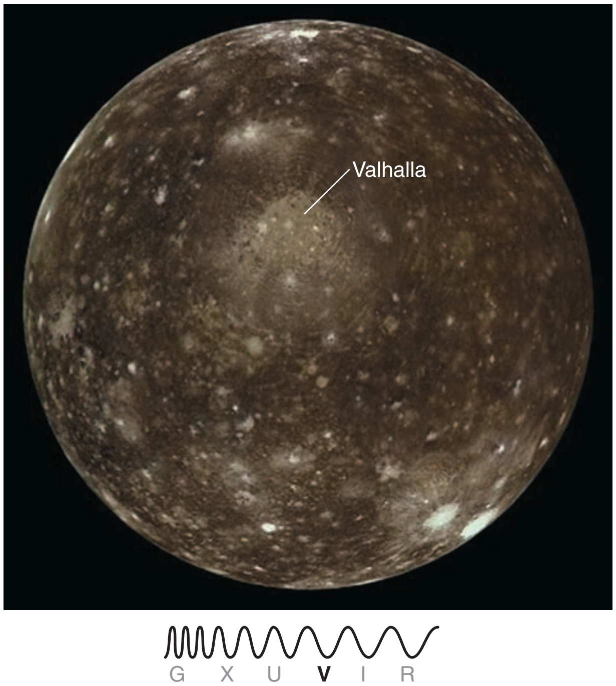


Figure 11.19 This *Galileo* image shows Jupiter's second-largest moon, Callisto. Its ancient surface is dominated by impact craters and shows no sign of early internal activity.

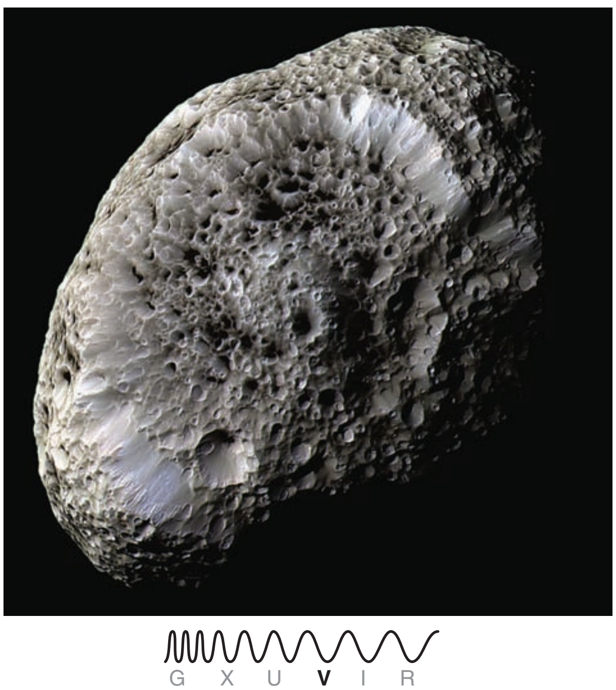


Figure 11.20 Saturn's moon Hyperion rotates chaotically, with its rotation period and spin axis constantly changing. The 250-km moon's low density and spongelike texture, seen in this *Cassini* image, suggest that its interior houses a vast system of caverns.

relatively dark, heavily cratered terrain. Callisto's most prominent feature is a 2,000-km, multiringed structure of impact origin named Valhalla (the largest bright feature visible on Callisto's face in **Figure 11.19**). *Galileo* results suggest that a liquid ocean containing water or water mixed with ammonia could exist beneath the heavily cratered surface. Callisto may have partially differentiated, with rocky material separating from ices and sinking deeper into the interior.

Saturn's Hyperion is one of the largest irregularly shaped moons and could be the remnant of an impact. The extensive craters look almost like sponges (**Figure 11.20**). Hyperion crosses Saturn's magnetosphere in its chaotic orbit, which seems to have left the moon with some electric charge. Umbriel, the darkest and third largest of Uranus's moons, appears uniform in color, reflectivity, and general surface features, indicative of an ancient surface. The real puzzle posed by Umbriel is why it is geologically dead, while the surrounding large moons of Uranus have been active at least at some time in their past.

CHECK YOUR UNDERSTANDING 11.3

Rank these moons in terms of the density of impact craters you would expect to observe on the surface. Rank them from most to least. (a) Callisto; (b) Titan; (c) Io; (d) Ganymede

11.3 Rings Surround the Giant Planets

A planetary ring is a collection of particles—varying in size from tiny grains to house-sized boulders—that orbit individually around a planet, forming a flat disk. Ring systems do not occur in the terrestrial planets but are found around each of the giant planets. **Figure 11.21** shows how the ring system of each giant planet varies in size and complexity: some systems extend for hundreds of thousands of kilometers, and some systems have detailed structure that includes numerous small rings. In this section, we discuss ring formation, composition, and evolution.

The Discovery of Planetary Rings

Saturn's rings have been observed for centuries. In 1610, Galileo observed two small objects next to Saturn and thought they might be similar to the four moons orbiting Jupiter. But Saturn's "moons" did not move, and 2 years later they disappeared. In 1655, Dutch instrument maker Christiaan Huygens (1629–1695) pointed a superior telescope of his own design at Saturn. Huygens observed that an apparently continuous flat ring surrounds the planet and that the ring's visibility changes with its apparent tilt as Saturn orbits the Sun. Over the next three centuries, astronomers discovered more rings around Saturn, but searches failed to detect rings around any other planet.

Most Solar System rings were more recently discovered. In 1977, a team of astronomers studying the atmosphere of Uranus during stellar occultations saw brief, minute changes in the brightness of a star as it first approached and then receded from the planet. The astronomers realized this meant that Uranus has rings. Over the next several years, stellar occultations revealed a total of nine rings surrounding the planet. In 1986, *Voyager 2* imaged two additional rings of

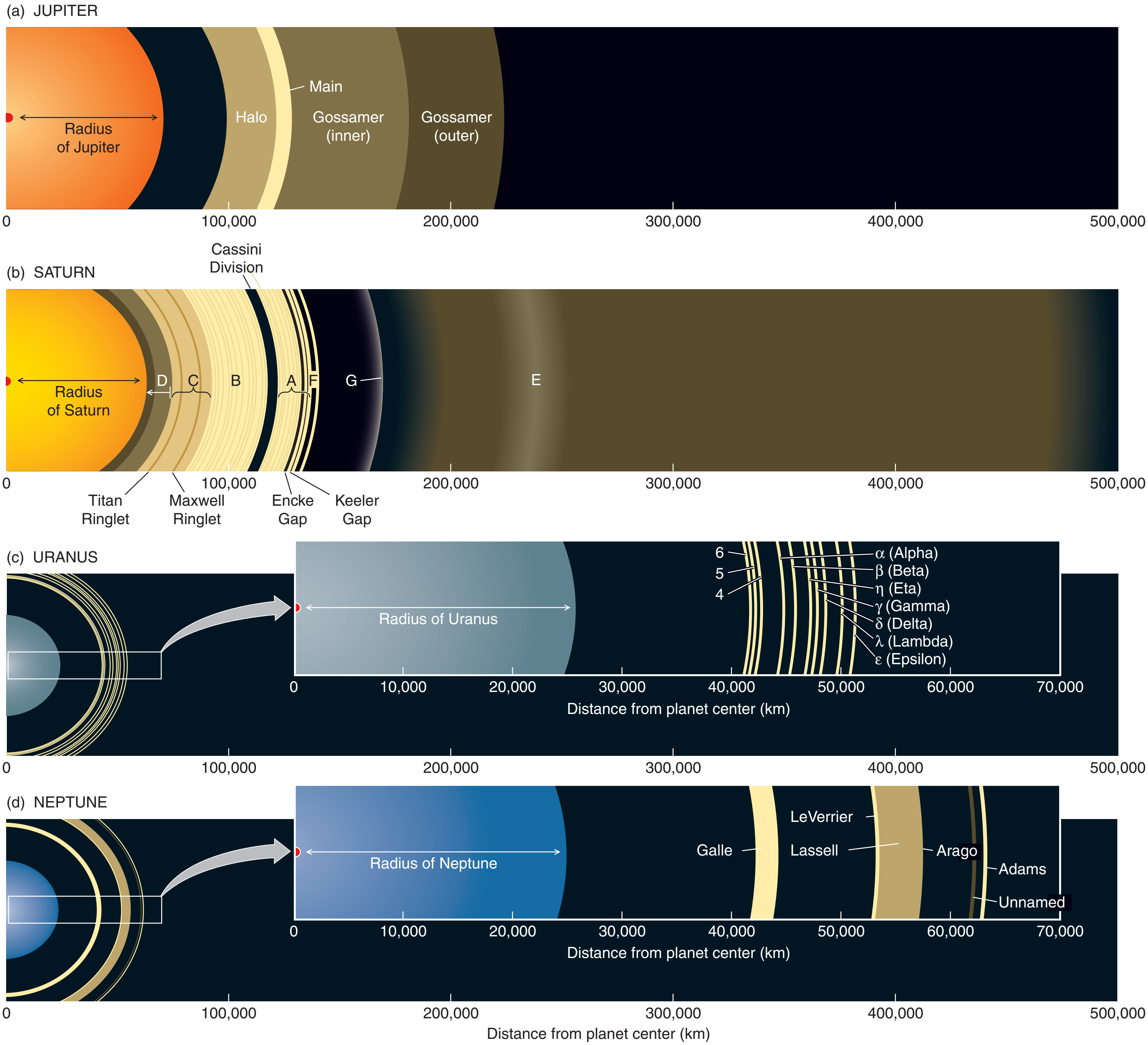


Figure 11.21 The ring systems of the four giant planets vary in size and complexity. Saturn's system, with its broad E Ring, is by far the largest and has the most complex structure in the inner rings.

Uranus, and in 2005 the Hubble Space Telescope recorded two more, bringing the total to 13. In 1979, cameras on *Voyager 1* recorded a faint ring around Jupiter. The occultation technique also revealed arclike ring segments around Neptune, which were determined to be complete rings when *Voyager 2* reached Neptune in 1989.

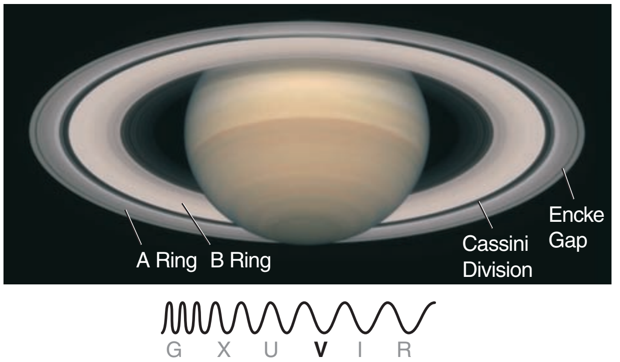


Figure 11.22 A Hubble Space Telescope image showing Saturn and its A Ring, B Ring, Cassini Division, and Encke Gap. The C Ring is too dim to be seen clearly.

The Orbits of Ring Particles

Ring particles follow Kepler's laws, and therefore the speed and orbital period of each particle must vary with its distance from the planet. The closest particles move the fastest and have the shortest orbital periods (see Working It Out 11.1). The orbital periods of particles in Saturn's bright rings, for example, range from 5 hours 45 minutes at the inner edge of the innermost bright ring to 14 hours 20 minutes at the outer edge of the outermost bright ring. Ring particles have low speeds relative to one another because they are all orbiting in the same direction. A particle moving on an upward trajectory will bump into another particle on a downward trajectory and the upward and downward motion will cancel, leaving the particles moving in the same plane. A similar process occurs for particles moving inward and outward, leaving the particles moving at a constant radius.

The orbits of ring particles can also be influenced by the planet's larger moons. If the moon is massive enough, it exerts a gravitational tug on the ring particles as it passes by. If this happens over and over through many orbits, the particles are pulled out of the area, leaving a lower-density gap (see Figure 11.21). Such is the case with Saturn's moon Mimas, which causes the famous gap in the rings around Saturn called the **Cassini Division** (**Figure 11.22**). Mimas is in a 1:2 orbital resonance with the Cassini Division, giving a ring particle located in the Cassini Division an orbital period about Saturn that is equal to half the orbital period of Mimas. Such resonances with other moons are known to produce some of the gaps that appear in Saturn's bright rings. One of the gaps is caused by a 4:1 resonance between the ring particles and Mimas.

Other kinds of orbital resonances are also possible. For example, most narrow rings are caught up in a periodic gravitational tug-of-war with nearby moons, known as shepherd moons because of the way they "herd" the flock of ring particles. Shepherd moons are usually small and often come in pairs, with one orbiting just inside and the other just outside a narrow ring. A shepherd moon just outside a ring robs orbital energy from any particles that drift outward beyond the edge of the ring, causing the particles to move back inward. A shepherd moon just inside a ring gives up orbital energy to a ring particle that has drifted too far in, nudging it back in line with the rest of the ring. In some cases, narrow rings are trapped between two shepherd moons in slightly different orbits.

Ring Formation and Evolution

Much of the material found in planetary rings is thought to be the result of tidal stresses. If a moon (or other planetesimal) orbits a large planet, the force of gravity will be stronger on the side of the moon close to the planet and weaker on the side farther away. This difference in gravitational force stretches out the moon, as you saw in the discussion of tidal forces in Chapter 4. If the tidal stresses are greater than the self-gravity that holds the moon together, the moon will be torn apart. The distance at which the tidal stresses exactly equal the self-gravity is known as the Roche limit. The Roche limit applies only to objects that are held together by their own gravity; it does not apply to objects held together by other forces—like people or cars. If a moon or planetesimal comes within the Roche limit of a planet, it is pulled apart by tidal stresses, leaving many small pieces orbiting the planet. These pieces gradually spread out, and their orbits are circularized and flattened out by collisions. The fragmented pieces of the disrupted body are then distributed around the planet in the form of a ring.

Planetary rings do not have the long-term stability of most Solar System objects. Ring particles are constantly colliding with one another in their tightly packed environment, either gaining or losing orbital energy. This redistribution of orbital energy can cause particles at the ring edges to leave the rings and drift away, aided by nongravitational influences such as the pressure of sunlight. Although moons may help guide the orbits of ring particles and delay the dissipation of the rings themselves, at best this condition can be only temporary. Saturn's brightest rings might be nearly as old as Saturn, but most planetary rings eventually disperse.

Even Earth may have had several short-lived rings at various times during its long history. Any number of comets or asteroids must have passed within Earth's Roche limit (about 25,000 km for rocky bodies and more than twice that for icy bodies) to disintegrate into a swarm of small fragments to create a temporary ring. However, unlike the giant planets, Earth lacks shepherd moons to provide orbital stability to rings.

The Composition of Ring Material

Because much of the material in the rings of the giant planets comes from their moons, the composition of the rings is similar to the composition of the moons. Saturn's bright rings probably formed when a moon or planetesimal came within the Roche limit of Saturn. These rings reflect about 60 percent of the sunlight falling on them. They are made of water ice, though a slight reddish tint tells us they are not made of pure ice but must contain small amounts of other materials, such as silicates. The icy moons around Saturn or the frozen comets of the outer Solar System could easily provide this material.

Saturn's rings are the brightest in the Solar System and are the only ones that we know are composed of water ice. In stark contrast, the rings of Uranus and Neptune are among the darkest objects known in the Solar System. Only 2 percent of the sunlight falling on them is reflected back into space, which makes the ring particles blacker than coal or soot. No silicates or similar rocky materials are this dark, so these rings are likely composed of organic materials and ices that have been radiation darkened by high-energy, charged particles in the magnetospheres of these planets. (Radiation blackens organic ices such as methane by releasing carbon from the ice molecules.) Jupiter's rings are of intermediate brightness, suggesting that they may be rich in silicate materials, like the innermost of Jupiter's small moons.

The jumble of fragments that make up Saturn's rings is understood to be a product of tidal disruption of a moon or planetesimal, but moons can contribute material to rings in other ways. The brightest of Jupiter's rings is a relatively narrow strand only 6,500 km across, consisting of material from the moons Metis and Adrastea. These two moons orbit in Jupiter's equatorial plane, and the ring they form is narrow. Beyond this main ring, however, are the very different wispy rings called gossamer rings. The gossamer rings are supplied with dust by the moons Amalthea and Thebe. The innermost ring in Jupiter's system, called the halo ring, consists mostly of material from the main ring. As the dust particles in the main ring drift slowly inward toward the planet, they pick up an electric charge and are pulled into this rather thick torus by electromagnetic forces associated with Jupiter's powerful magnetic field.

Finally, moons may contribute ring material through volcanism. Volcanoes on Jupiter's moon Io continually eject sulfur particles into space, many of which are

11.3 Working It Out Feeding the Rings

The moons of the giant planets have a low surface gravity and a much lower escape velocity than that of Earth, which is 11.2 km/s. Thus, volcanic emissions from some of these small moons can escape and supply material to a ring. Recall the equation from Working It Out 4.2 for escape velocity from a spherical object of mass M and radius R:

$$v_{esc} = \sqrt{\frac{2GM}{R}}$$

Saturn's moon Enceladus has a mass of 1.08×10^{20} kg and a radius of 250 km. The escape velocity from Enceladus is given by

$$v_{esc} = \sqrt{\frac{2 \times [6.67 \times 10^{-20}\,\text{km}^3/(\text{kg s}^2)] \times (1.08 \times 10^{20}\,\text{kg})}{250\,\text{km}}}$$

$$v_{esc} = 0.24\,\text{km/s; or multiply by 3,600 s/h to get 864 km/h}$$

This escape velocity is lower than the speed of the volcanic plumes on Enceladus, which is nearly 2,200 km/h. The icy particles from the plumes supply particles to Saturn's E Ring.

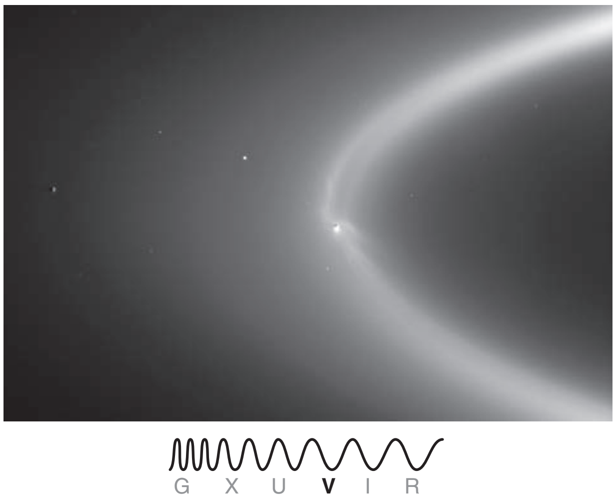


Figure 11.23 Saturn's moon Enceladus (the large bright spot appearing to be on the ring) is the source of material in Saturn's E Ring. Note the distortion in the distribution of ring material in the immediate vicinity of the moon caused by its gravitational influence on the orbits of ring particles. Other bright objects in this *Cassini* image are stars.

pushed inward by sunlight and find their way into a ring. The particles in Saturn's E Ring are ice crystals ejected from icy geysers on the moon Enceladus, which is located in the very densest part of the E Ring (**Working It Out 11.3**). Ice particles ejected into space replace particles continually lost from Saturn's E Ring (**Figure 11.23**). The E Ring will survive for as long as Enceladus remains geologically active.

CHECK YOUR UNDERSTANDING 11.4

If rings are observed around a planet, this indicates that: (a) there is a recent source of ring material; (b) the planet is newly formed; (c) the rings formed with the planet; (d) the rings are made of fine dust.

11.4 Ring Systems Have a Complex Structure

Huygens, with his mid-17th-century understanding of physics, thought that Saturn's ring was a solid disk surrounding the planet. It was not until the middle of the 19th century that the brilliant Scottish mathematical physicist James Clerk Maxwell showed that solid rings would be unstable and would quickly break apart. In this section, we will examine the details of the rings of each planet of the outer planets. You may want to refer back to Figure 11.21.

Saturn's Magnificent Rings—A Closer Look

Saturn is adorned by a magnificent and complex system of rings, unmatched by any other planet in the Solar System. Figure 11.21b shows the individual components of Saturn's ring system and its major divisions and gaps. Among the four giant planets, Saturn's rings are the widest and brightest. The outermost bright ring, the A Ring, is the narrowest of the three bright rings. It has a sharp outer edge and contains several narrow gaps.

In 1675, the Italian-French astronomer Jean-Dominique Cassini (1625–1712) found a gap in the planet's seemingly solid ring. Saturn appeared to have two rings rather than one, and the gap that separated them became known as the *Cassini Division*. The Cassini Division is so wide (4,700 km) that the planet Mercury would almost fit within it. Astronomers once thought that it was completely empty, but images taken by *Voyager 1* show the Cassini Division is filled with material, although it is less dense than the material in the bright rings.

The B Ring, whose width is roughly twice Earth's diameter, is the brightest of Saturn's rings and has no internal gaps on the scale of those seen in the other bright rings. The C Ring is so much fainter than neighboring rings that it often fails to show up in normally exposed photographs. Through the eyepiece of a telescope, this ring appears like delicate gauze. There is no known gap between the C Ring and either of the adjacent rings; only an abrupt change in brightness marks the boundary between them. The cause of this sharp change in the amount of ring material remains an unanswered question. Too dim to be seen next to Saturn's bright disk, the D Ring is a fourth wide ring that was unknown until it was imaged by *Voyager 1*. It shows less structure than any of the bright rings, and it does not appear to have a definable inner edge. The D Ring may extend all the way down to the top of Saturn's atmosphere, where its ring particles would burn up as meteors.

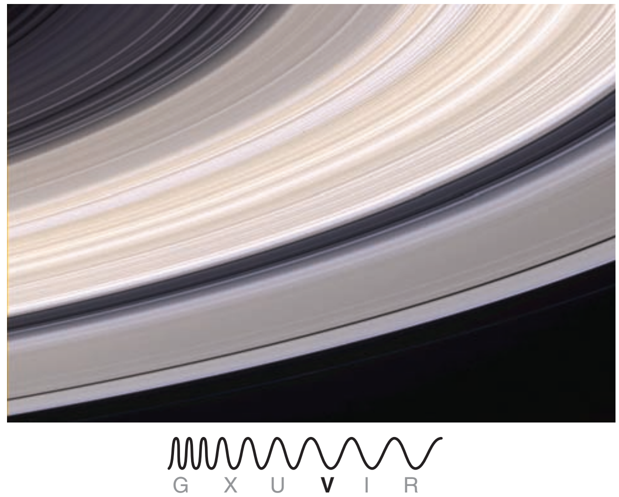
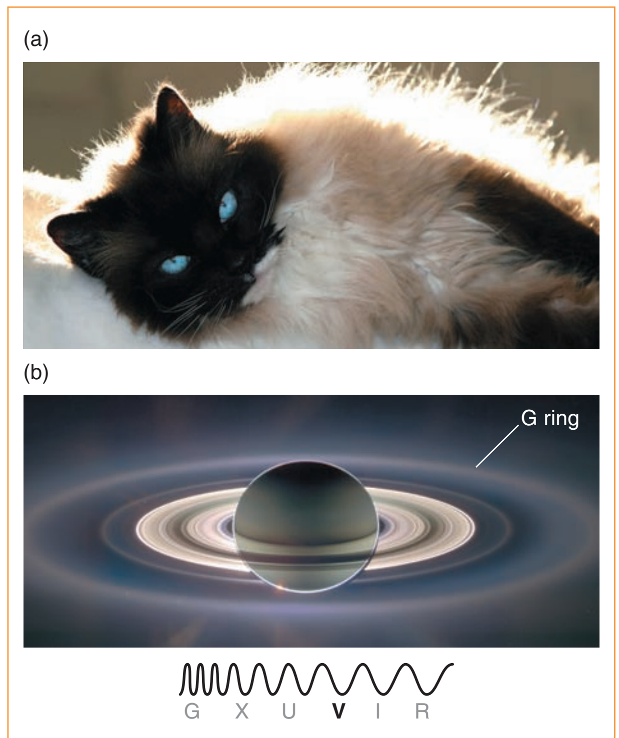


Figure 11.24 This *Cassini* image of the rings of Saturn shows so many ringlets and minigaps that it looks like a close-up of an old-fashioned phonograph record. The cause of most of this structure has yet to be explained in detail.

Saturn's bright rings are not uniform. The A and C rings contain hundreds, and the B Ring thousands, of individual ringlets, some only a few kilometers wide (**Figure 11.24**). Each of these ringlets is a narrowly confined concentration of ring particles bounded on both sides by regions of relatively little material. About every 15 years, the plane of Saturn's rings lines up with Earth, and we view them edge on. The rings are so thin that they all but vanish for a day or so in even the largest telescopes. While the glare of the rings is absent, astronomers search for undiscovered moons or other faint objects close to Saturn. In 1966, an astronomer was looking for moons when he found weak but compelling evidence for a faint ring near the orbit of Saturn's moon Enceladus. In 1980, *Voyager 1* confirmed the existence of this faint ring, now called the E Ring, and found another closer one known as the G Ring.

The E and G rings are examples of diffuse rings. In a diffuse ring, particles are far apart, and rare collisions between them can cause their individual orbits to become eccentric, inclined, or both. Because collisions are rare, the particles tend to remain in these disturbed orbits. Diffuse rings spread out horizontally and thicken vertically, sometimes without any obvious boundaries.

Diffuse rings contain tiny particles that show up best when the viewer is looking at these rings in the direction of the Sun. In contrast, larger objects such as pebbles and boulders are easiest to see when the light illuminating them is coming from behind the viewer. Dust and other small particles, however, stand out most strongly when you look *into* the light. For example, dust particles on your windshield appear brightest when you are driving toward the Sun. Photographers call this effect backlighting and often place their subjects in front of a bright light to highlight hair (**Figure 11.25a**). Backlighting happens when light falls on very small objects—those with dimensions a few times to several dozen times the wavelength of light. Cat fur and human hair is near the upper end of this range. Light falling on strands of hair tends to continue in the direction away from the source of illumination. Very little of the light is scattered off to the side, and almost none is scattered back toward the source.



VISUAL ANALOGY

Figure 11.25 (a) Backlighting of hair creates a halo effect. (b) This backlit image of Saturn shows the G Ring.

Some of the dustier planetary rings are filled with particles that are just a few times larger than the wavelength of visible light. To a spacecraft approaching from the direction of the Sun, such rings may be difficult or even impossible to

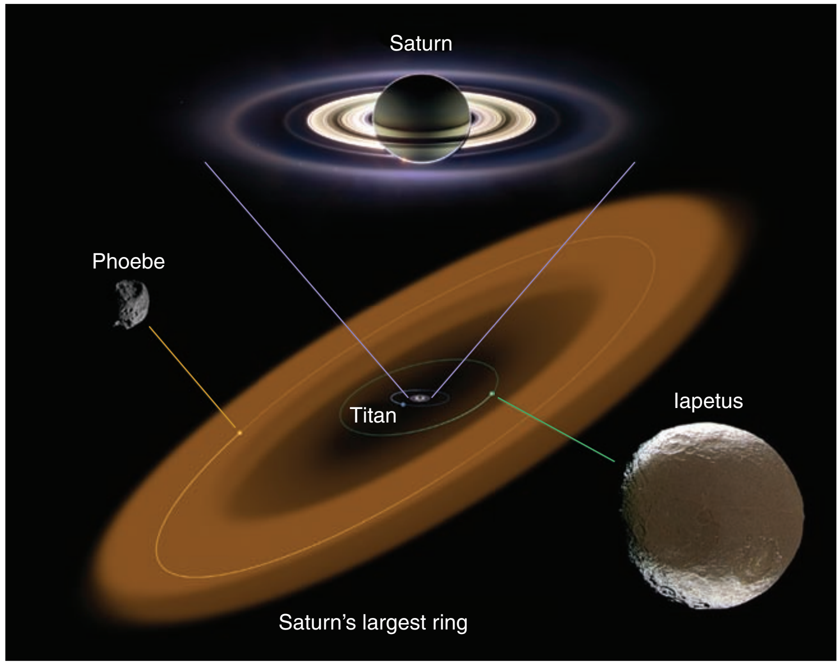


Figure 11.26 This artist's conception shows the highly inclined giant dust ring recently discovered around Saturn. This ring is so large that the rest of Saturn appears as a speck in the center (magnified in the inset).

see. These tiny ring particles scatter very little sunlight back toward the Sun and the approaching spacecraft. However, when the spacecraft passes by the planet and looks backward in the general direction of the Sun, these same dusty rings suddenly appear as a circular blaze of light, much like a halo surrounding the nighttime hemisphere of the planet. Many planetary rings are best seen with backlighting, and some, such as Saturn's G Ring (see Figure 11.25b), have been observed only under these conditions. In 2009, astronomers using the infrared Spitzer Space Telescope discovered another diffuse ring around Saturn. This dusty ring is thicker than other rings, about 20 times larger than Saturn from top to bottom, and is tilted 27° with respect to the plane of the rest of the rings (**Figure 11.26**).

Although Saturn's bright rings are very wide—more than 62,000 km from the inner edge of the C Ring to the outer edge of the A Ring—they are extremely thin. Saturn's bright rings are no more than 100 meters thick and probably only a few tens of meters from their lower to upper surfaces. The diameter of Saturn's bright ring system is 10 million times the thickness of the rings. If the bright rings of Saturn were the thickness of a page in a book, six football fields laid end to end would stretch across them.

Voyager 1 images showed that Saturn's F Ring is separated into several strands that appear to be intertwined and also display what appear to be a number of knots and kinks. Saturn's F Ring is now understood to be a dramatic example of the action of a pair of shepherd moons. The F Ring is flanked by Prometheus, a moon that orbits 860 km inside the ring, and Pandora, a moon that orbits 1,490 km on the outside, as seen in a more recent *Cassini* image (**Figure 11.27**). Both moons are irregular in shape, with average diameters of 85 and 80 km, respectively. Because of their relatively large size and proximity, the moons exert significant gravitational forces on nearby ring particles. The resulting tug-of-war between Prometheus pulling ring particles in its vicinity into larger orbits and Pandora drawing its neighboring particles into smaller orbits is the cause of the bizarre structure in the F Ring (**Process of Science Figure**).

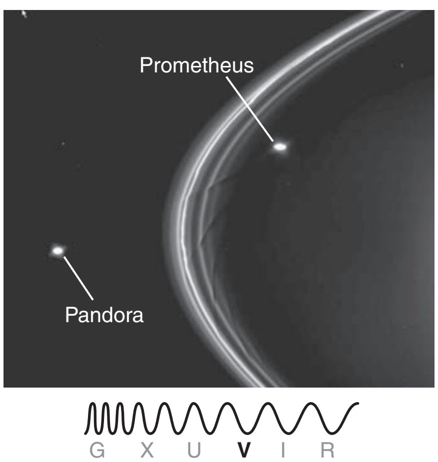


Figure 11.27 This *Cassini* image shows Saturn's F Ring and its shepherd moons Pandora and Prometheus. You can also see some "kinks" in the inner ring.

Process of Science

FOLLOWING UP ON THE UNEXPECTED

Scientists expected to find the dust particles in the rings of Saturn moving on undisturbed orbits. Instead, the F Ring particles seemed to disobey the laws of physics!

The *Voyager 1* spacecraft discovers multiple intertwining strands, knots, and kinks in Saturn's F Ring.

Scientists proclaim: "We must not have accounted for everything."

The media proclaim: "The laws of physics are wrong!"

Further observations reveal previously unobserved moons that produce the unexpected behavior of F Ring particles.

Theorists verify gravitational effects with simulations.

Observers find more examples of deformed rings and the "shepherd moons" that cause them.

The theory that rings can be distorted by the gravitational influence of nearby moons becomes widely accepted as it repeatedly passes tests of observation and simulation.

Scientists are excited by apparent violations of well-supported theories because it may lead to a new discovery. One consequence is that unexpected or contradictory results often receive more attention than confirming results.

Figure 11.28 In this *Cassini* high-resolution view, Saturn's Encke Gap reveals a scalloped pattern along its inner edge that is caused by the moon Pan.

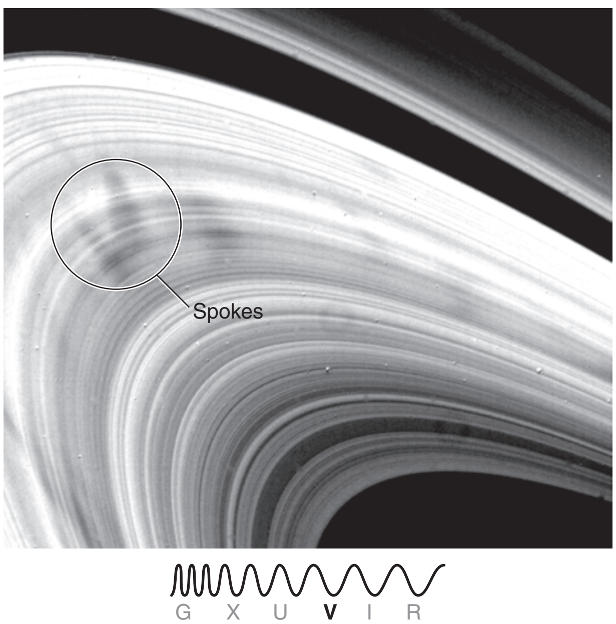


Figure 11.29 Spokes in Saturn's B Ring appear dark in normal viewing but bright with backlighting.

The F Ring is not an isolated case. The 360-km-wide Encke Gap in the outer part of Saturn's A Ring contains two narrow rings that show bright knots and dark gaps—a structure that must be related to a 20-km-diameter moon named Pan that orbits within the gap. Small moons orbiting within ring gaps can also disturb ring particles along the edges of the gaps. **Figure 11.28** shows the scalloped pattern caused by Pan that is found along the inner edge of the Encke Gap. Similarly, the 7-km-diameter moon Daphnis disrupts the inner and outer edges of Saturn's 35-km-wide Keeler Gap, located near the outer edge of the A Ring.

Voyager 1 and then *Cassini* observed dozens of dark, spokelike features in the outer part of Saturn's B Ring (**Figure 11.29**). These temporary features grow in a radial direction and last for less than half an orbit around Saturn, indicating that the particles in the spokes must be suspended above the ring plane, probably by electrostatic forces. One explanation is that when the charged particles interact with Saturn's magnetic field, the spokes rotate as the planet spins.

Rings around the Other Outer Planets

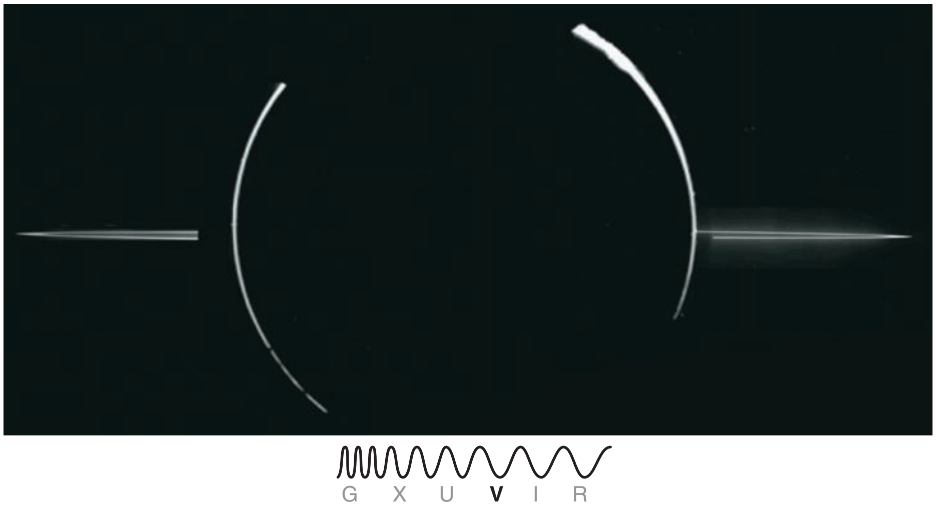
Ring structure among the other giant planets is not as diverse as Saturn's. Most rings other than Saturn's are quite narrow, although a few are diffuse. When *Voyager 1* scientists looked at Jupiter's ring system with the Sun behind the camera, all they saw was a narrow, faint strand. But when *Voyager 2* looked back toward the Sun while in the shadow of the planet, Jupiter's rings suddenly blazed into prominence. **Figure 11.30a** shows a nearly edge-on, backlit view of Jupiter's rings taken by the *Galileo* spacecraft. Most of the material in Jupiter's rings is made up of fine dust dislodged by meteoritic impacts on the surfaces of Jupiter's small inner moons. The moons are shown orbiting among the rings in Figure 11.30b.

Of the 13 rings of Uranus, 9 are very narrow and widely spaced relative to their widths (see Figure 11.21c). Most are only a few kilometers wide, but they are many hundreds of kilometers apart. The two rings discovered by the Hubble Space Telescope in 2005 (**Figure 11.31**) are much wider and more distant than the narrow rings. The most prominent ring of Uranus, the Epsilon Ring, is eccentric and the widest of the planet's inner narrow rings, varying in width between 20 and 100 km. The innermost ring is wide and diffuse, with an undefined inner edge. As with Saturn's D Ring, material in this ring may be spiraling into the top of the planetary atmosphere. When viewed under backlit conditions by *Voyager 2*, the space *between* the rings of Uranus turned out to be filled with dust, much as in Jupiter's ring system.

The rings in the Encke Gap are unusual but not unique. If shepherd moons are in eccentric or inclined orbits, they cause the confined ring also to be eccentric or inclined. This is the case for the Epsilon Ring of Uranus. Because shepherd moons can be so small, they often escape detection. According to current theories of ring dynamics, a number of still-unknown shepherd moons must be interspersed among the ring systems of the outer Solar System.

For a while, Neptune seemed to be the only giant planet devoid of rings. Then, in the early to mid 1980s, occultation searches by teams of astronomers began yielding confusing results. Several occultation events that appeared to be due to rings were seen on only one side of the planet. The astronomers concluded that Neptune was surrounded not by complete rings but rather by several arclike ring

(a)



(b)

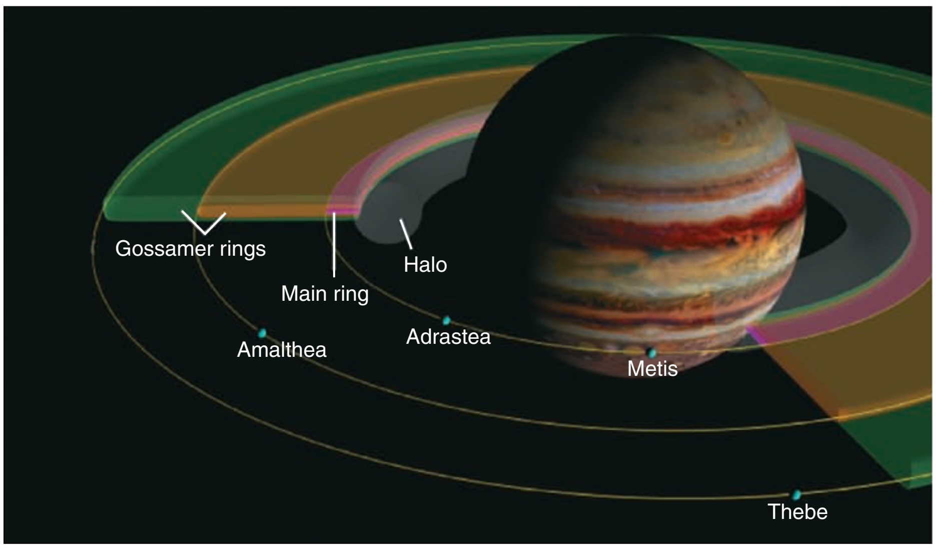


Figure 11.30 (a) This backlit *Galileo* image of Jupiter's rings also shows the forward scattering of sunlight by tiny particles in the upper layers of Jupiter's atmosphere. (b) A diagram of the Jupiter ring system and the small moons that form the rings.

segments. When *Voyager 2* reached Neptune in 1989 it was determined that Neptune's rings are complete. The **ring arcs** are high-density segments within one of its narrow rings. All of Neptune's rings are faint and, with the exception of the ring arcs, they contain too little material to be detected by the stellar occultation technique.

Four of Neptune's six rings are very narrow, similar to the 13 narrow rings surrounding Uranus. The other two have widths of a few thousand kilometers (see Figure 11.21d). Neptune's rings are named for 19th century astronomers who made major contributions to Neptune's discovery. Of these, the Adams Ring attracts the greatest attention. Much of the material in the Adams Ring is clumped together into several ring arcs. These high-density ring segments extend over lengths of 4,000–10,000 km, yet are only about 15 km wide. When first discovered, the ring arcs were a puzzle, because mutual collisions among their particles

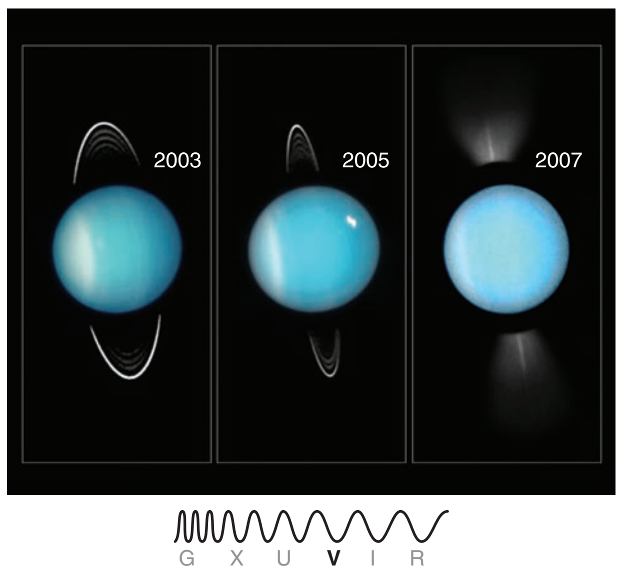


Figure 11.31 The appearance of rings depends dramatically on lighting conditions and the angle from which they are seen. Earth's view of the rings of Uranus changes over the course of several years, as shown here.

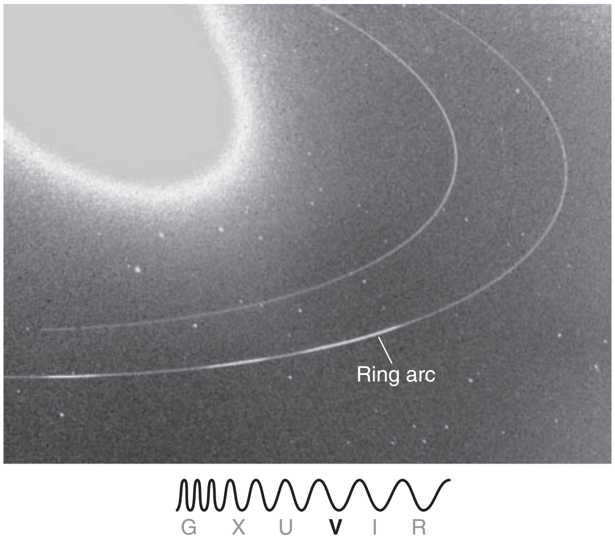


Figure 11.32 This *Voyager 2* image shows the three brightest arcs in Neptune's Adams Ring. Neptune itself is very much overexposed in this image.

should cause the particles to be spread more or less uniformly around their orbits. Most astronomers now attribute this clumping to orbital resonances with the moon Galatea that orbits just inside the Adams Ring (**Figure 11.32**). Images obtained by the Hubble Space Telescope in 2004 and 2005 compared with those taken by *Voyager 2* in 1989 show that some parts of the Neptune ring arcs are unstable. Slow decay is seen in two of the arcs, Liberté and Courage, suggesting that they may disappear entirely before the end of the century. Uranus's Lambda Ring and Saturn's G Ring also show ring arcs.

Moons and Rings around Extrasolar Planets

In previous chapters, we have shown extrasolar planets are common. Large and small planets, many in multiplanet systems, have been detected within our galaxy. If our Solar System is typical, then we might expect that other planetary systems contain planets with rings and planets with large moons, perhaps with geological activity and water. However, the identification of extrasolar moons and rings is at the limit of current astronomical instrumentation. As of this writing, there have not been confirmed detections.

The proposed methods for detecting *exomoons* are similar to those for detecting exoplanets. Recall from Chapter 7 that *Kepler* detects planets through the transit method (see Figures 7.19 and 7.20). The presence of a moon near the planet can slightly alter the depth and duration of the change in the light curve. The moon is likely in a different place in its orbit each time the planet orbits the star, so these alterations in the light curve will be different for each cycle. Over the course of several cycles, the signature of a moon might be detected. It is also possible that astronomers could detect a large exomoon as the moon itself transits its star. Or a large moon could make its planet "wobble" in its orbit, and this wobble could be detected in the transit signal. NASA supercomputers are being used to analyze the large database from *Kepler* to look for signatures of exomoons.

Similarly, a large extrasolar planet with an extensive ring system, especially if the ring system has gaps, might be detectable through changes in the transit signal. The depth and length of the changes in the star's light curve could indicate the presence of such a system. **Figure 11.33** is an artist's conception of an extrasolar planet with rings and a large exomoon.

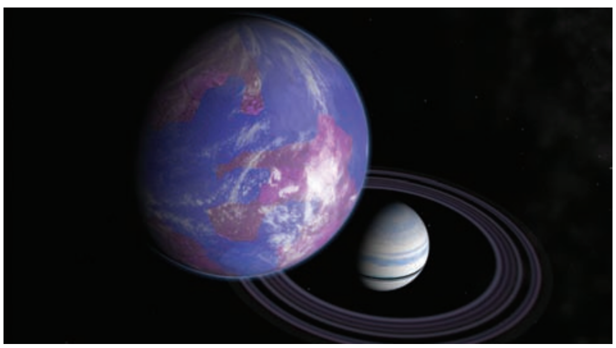


Figure 11.33 This artist's conception shows a hypothetical Earth-like moon around a Saturn-like exoplanet.

CHECK YOUR UNDERSTANDING 11.5

If you wanted to search for faint rings around a giant planet by sending a spacecraft on a flyby, it would be best to make your observations: (a) as the spacecraft approached the planet; (b) after the spacecraft passed the planet; (c) while orbiting the planet; (d) during the closest flyby; (e) while orbiting one of its moons.

Origins

Extreme Environments

During the 1980s and 1990s, as the *Voyager* spacecraft were exploring the outer Solar System, biologists back on Earth were identifying strange forms of life. Off the coast of the Galápagos Islands, 2,500 meters beneath the ocean's surface, plates grind against one another, creating friction, high temperatures, and seafloor volcanism. Mineral-rich, superheated water pours out of hydrothermal vents. The surrounding water contains very little dissolved oxygen. No sunlight reaches these depths, yet in the total darkness of the ocean bottom, life abounds. From tiny bacteria to shrimp to giant clams and tube worms, sea life thrives in this severe environment. In the complete absence of sunlight, the small, single-celled organisms at the bottom of the local food chain get their energy from *chemosynthesis*, a process by which inorganic materials are converted into food through the use of chemical energy. Biologists call these life-forms *extremophiles*.

Similarly, robust types of bacteria are found flourishing in the scalding waters of Yellowstone's hot springs; in the bone-dry oxidizing environment of Chile's Atacama Desert; and in the Dead Sea, where salt concentrations run as high as 33 percent. Bacteria have even been found in core samples of ancient ice 3,600 meters below the surface of the East Antarctic ice sheet. When it comes to harsh habitats, life is amazingly adaptable. If life can exist under such extreme conditions on Earth, might it also exist on those moons of the giant planets that have the ingredients necessary for life on Earth: liquid water, an energy source, and the presence of organic compounds?

When scientists realized that Mars and Venus were not as Earth-like as once imagined, prospects for finding life elsewhere in the Solar System seemed dim. Now astrobiologists are turning their attention to some of the small worlds that circle the giant planets far from the Sun. These moons may supply clues about the history of life in the Solar System. Their environments may be similar to some of the ecological niches on Earth that support extremophiles. The conditions necessary to create and support life on Earth—liquid water, heat, and organic material—could all be present in oceans on some of these moons.

Enceladus, Saturn's geologically active ice moon, spews salty water-ice grains, indicating liquid water below the surface. Perhaps its south polar region is a habitable zone. The fractured ice floes that make up the surface of Jupiter's Europa may cover an ocean warmed and enriched by geothermal vents not unlike those that dot the floors of Earth's oceans. Like Europa, Callisto also shows magnetic variability, a possible signature of a salty ocean. Saturn's Titan has an atmosphere and a possible subsurface ocean, and it is similar in several ways to a much earlier Earth.

The presence of comet-borne organic material in large bodies of water on Europa, Callisto, Enceladus, or Titan cannot yet be confirmed. Methane can arise from biological processes or it can come from chemical or geochemical processes, so its detection on Enceladus and Titan is tantalizing but is not evidence of life. In addition to the presence of methane, spectroscopy reveals organic gases in Titan's massive atmosphere. Titan's nitrogen atmosphere contains compounds of biological interest. For example, five molecules of hydrogen cyanide (HCN) will spontaneously combine to form adenine, one of the four primary components of DNA and RNA. HCN is also a building block of amino acids, which in turn combine to form proteins. Photodissociation and recombination of these various gases produce complex organic molecules that then rain out onto Titan's surface as a frozen tarry sludge. Biochemists think that many of these substances are biological precursors, similar to the organic molecules that preceded the development of life on Earth.

Astronomers anticipate future exploration of these moons, which may provide fascinating clues to the origins of terrestrial life.

ARTICLES QUESTIONS

Astronomers report that a new moon of Saturn might be forming.

Possible New Moon Forming around Saturn

By **Science@NASA**

NASA's *Cassini* spacecraft has documented the formation of a small icy object within the rings of Saturn. Informally named "Peggy," the object may be a new moon. Details of the observations were published online today by the journal *Icarus*.

"We have not seen anything like this before," said Carl Murray of Queen Mary University of London, and the report's lead author. "We may be looking at the act of birth, where this object is just leaving the rings and heading off to be a moon in its own right."

Images taken with *Cassini*'s narrow-angle camera on April 15, 2013, show disturbances at the very edge of Saturn's A Ring—the outermost of the planet's large, bright rings (**Figure 11.34**). One of these disturbances is an arc about 20 percent brighter than its surroundings, 750 miles (1,200 kilometers) long and 6 miles (10 kilometers) wide. Scientists also found unusual protuberances in the usually smooth profile at the ring's edge. Scientists believe the arc and protuberances are caused by the gravitational effects of a nearby object.

The object is not expected to grow any larger, and may even be falling apart. But the process of its formation and outward movement aids in our understanding of how Saturn's icy moons, including the cloud-wrapped Titan and ocean-holding Enceladus, may have formed in more massive rings long ago. It also provides insight into how Earth and other planets in our Solar System may have formed and migrated away from our star, the Sun.

"Witnessing the possible birth of a tiny moon is an exciting, unexpected event," said *Cassini* Project Scientist Linda Spilker, of NASA's Jet Propulsion Laboratory (JPL) in Pasadena, California. According to Spilker, *Cassini*'s orbit will move closer to the outer edge of the A Ring in late 2016 and provide an opportunity to study Peggy in more detail and perhaps even image it.

Figure 11.34 The small object "Peggy" is at the edge of the ring.

Peggy is too small to see in images so far. Scientists estimate it is probably no more than about a half mile in diameter. Saturn's icy moons range in size depending on their proximity to the planet—the farther from the planet, the larger. And many of Saturn's moons are composed primarily of ice, as are the particles that form Saturn's rings. Based on these facts, and other indicators, researchers recently proposed that the icy moons formed from ring particles and then moved outward, away from the planet, merging with other moons on the way.

"The theory holds that Saturn long ago had a much more massive ring system capable of giving birth to larger moons," Murray said. "As the moons formed near the edge, they depleted the rings."

It is possible the process of moon formation in Saturn's rings has ended with Peggy, as Saturn's rings now are, in all likelihood, too depleted to make more moons. Because they may not observe this process again, Murray and his colleagues are wringing from the observations all they can learn.

ARTICLES QUESTIONS

1. What does this article suggest about the link between the formation of moons and rings?
2. Is a moon this size likely to be geologically active? Explain.
3. Some, but not all, of Saturn's moons are listed in Appendix 4 (and on the websites in end-of-chapter question 47). Is it correctly stated that the farther the icy moon is from the planet, the larger it is in size?
4. Looking at the names of the other moons of Saturn, is this moon likely to keep the name "Peggy"?
5. Has *Cassini* revisited this area of the rings? Do a search to see what it found.

Summary

The moons of the outer Solar System are composed of rock and ice. A few moons are geologically active, but most are dead. All four giant planets have ring systems, which are transitory, created from and maintained by moons also in orbit around these planets. Scientists are excited about the evidence for oceans on several of the moons and the possibility that some form of life may exist in these oceans.

LG 1 **Compare and contrast the formation and orbits of regular and irregular moons.** Most of the regular moons were formed along with their parent planets and have short and nearly circular orbits. Irregular moons were captured later, have more elongated orbits, and often travel in the opposite direction of the planet's rotation. Observations of the orbits of moons can be used to find the masses of their host planets.

LG 2 **Describe the evidence for geological activity and liquid oceans on some of the moons.** Jupiter's Io is the most volcanically active body in the Solar System. Jupiter's moon Europa contains an enormous subsurface ocean and probably has some geological activity. Saturn's moon Titan has lakes of liquid methane and perhaps a deep, salty ocean. Saturn's moon Enceladus and Neptune's moon Triton have cryovolcanoes. The large Galilean moons of Jupiter, Ganymede and Callisto, may also have subsurface oceans. Some moons were geologically active in the past, as indicated by crater scars and areas smoothed by flowing fluids. Moons that have always been geologically dead show nothing but impact craters on their surfaces.

LG 3 **Describe the composition, origin, and general structure of the rings of the giant planets.** Rings are formed of countless numbers of particles all in the same plane, held to the host planet by gravity. Some rings form when moons cross a planet's Roche limit. The composition of these moons determines the composition of the rings that form from them. Shepherd moons often maintain and shape rings by gravitationally pulling and pushing these particles as they pass by. Ring particles also interact gravitationally with each other. Some rings may be transient features held in place by moons. Saturn's bright rings and its E Ring are made primarily of water ice: the rings of the other planets are composed of darker materials.

LG 4 **Explain the role gravity plays in the structure of the rings and the behavior of ring particles.** Saturn's ring system is the most complex, and it is the best laboratory for understanding gravitational interactions between moons and rings, interactions between rings, and ring formation and dissipation. Gravity holds the ring particles in orbit around the planet, and gravitational interactions with moons determines the size and shape of rings.

? UNANSWERED QUESTIONS

- What is the source of Titan's nitrogen atmosphere? One group of experimenters studied this by blasting a laser at water-ammonia (H_2O-NH_3) ice to simulate cometary impacts and see whether nitrogen gas (N_2) forms. They concluded that the observed amount of N_2 in Titan's atmosphere could have been created from ammonia ice in this way. Another theory is that these gases were accreted during the formation of the moon. Using data from the *Huygens* probe, other researchers conclude that if Titan had differentiated like Ganymede, then hydrothermal activity released the gases from a hot core. Astronomers want to understand why Titan has an atmosphere and the other larger moons do not.
- What is the origin of Saturn's brightest rings? One early hypothesis is that the rings come from a moon that approached too close to the planet. But Saturn's moons are composed of rock and ice, and the rings are solely ice. Some recent computer models start with a differentiated moon the size of Titan that had a rock and iron core and a large, icy mantle. As the moon in the model slowly migrates toward Saturn and crosses the Roche limit, tidal forces rip away its water ice, but not the rocky core. According to this model, the core might have continued migrating inward until it fell into Saturn, and the ice would have formed Saturn's ring. As time went on, the ring spread, and as material crossed the Roche limit outward, Saturn's small moonlets formed. Computers are only now getting fast enough fully to test these models, which suggest a unified origin for many of the moons and rings.
- When will there be robotic space missions to the outer moons? Several missions have been proposed to study Europa, Ganymede, Titan, and Enceladus. Missions to Europa to investigate the liquid ocean beneath the icy surface have the highest priority. One proposed mission would crash a probe onto the surface of Europa and collect and analyze the debris plume. Another would have a probe land on the ice and drill into it until it reached liquid water. (This has been done at Lake Vostok, an under-ice lake in Antarctica.) Proposed missions to Titan include a balloon that would hover in its atmosphere for many months taking data or a probe that would land and float on one of the lakes. NASA and the European Space Agency are discussing possible missions for launch in the mid-2020s.

Questions and Problems

Test Your Understanding

1. Categorizing moons by geological activity is helpful because
   a. comparing them reveals underlying physical processes.
   b. geological activity levels drop with distance from the Sun.
   c. geological activity determines the size and composition of the moons.
   d. most moons are very similar to each other.

2. Why are Ganymede and Callisto geologically dead while the other two Galilean moons of Jupiter are active?
   a. they are larger
   b. they are farther from Jupiter
   c. they are more massive
   d. they have retrograde orbits

3. Moons of outer planets may provide a home for life because
   a. some have liquid water.
   b. some have organic molecules.
   c. some have an interior source of energy.
   d. all of the above

4. Io has the most volcanic activity in the Solar System because
   a. it is continually being bombarded with material in Saturn's E Ring.
   b. it is one of the largest moons and its interior is heated by radioactive decays.
   c. of gravitational friction caused by the moon Enceladus.
   d. its interior is tidally heated as it orbits around Jupiter.
   e. the ice on the surface creates a large pressure on the water below.

5. Gravitational interactions with moons produce
   a. fine structure within rings.
   b. short-lived rings.
   c. smoothed-out rings.
   d. rings with spokes.

6. Saturn's bright rings are located within the Roche limit of Saturn. This fact supports the theory that these rings (select all that apply)
   a. formed of moons torn apart by tidal stresses.
   b. formed at the same time that Saturn formed.
   c. are relatively recent.
   d. are temporary.

7. The story of the F Ring of Saturn is an example of
   a. an unexplained phenomenon.
   b. media bias.
   c. the self-correcting nature of science.
   d. a violation of causality.

8. If a moon revolves opposite to its planet's rotation, it probably
   a. was captured after the planet formed.
   b. had its orbit altered by a collision.
   c. has a different composition from other moons.
   d. formed very recently in the Solar System's history.

9. Under what lighting conditions are the tiny dust particles found in some planetary rings best observed?
   a. viewed from the shadowed side of the planet, looking toward deep space
   b. viewed from the shadowed side of the planet, looking toward the Sun
   c. viewed from the near side of the planet, looking toward deep space
   d. viewed from the near side of the planet, looking toward the Sun

10. Planets in the outer Solar System have more moons than those in the inner Solar System because
    a. the solar wind was weaker there.
    b. there was more debris around the outer planets when they were forming.
    c. the outer planets captured most of their moons.
    d. there were more planetesimal collisions far from the Sun.

11. The difference between a moon and a planet is that
    a. moons orbit planets, whereas planets orbit stars.
    b. moons are smaller than planets.
    c. moons and planets have different compositions.
    d. moons and planets formed in different ways.

12. Scientists determine the geological history of the moons of the outer planets from
    a. seismic probing.
    b. radioactive dating.
    c. surface features.
    d. time-lapse photography.

13. The energy that keeps Io's core molten comes from
    a. the Sun.
    b. radioactivity in the core.
    c. residual heat from the collapse.
    d. Jupiter's gravity.

14. We classify moons as formerly active if they
    a. are completely covered in craters.
    b. have no young craters.
    c. have regions with few craters.
    d. have regions with no craters.

15. Volcanoes on Enceladus affect the E Ring of Saturn by
    a. pushing the ring around.
    b. stirring the ring particles.
    c. supplying ring particles.
    d. dissipating the ring.

Thinking about the Concepts

16. Explain the process that drives volcanism on Jupiter's moon Io.
17. Describe cryovolcanism and explain its similarities and differences with respect to terrestrial volcanism. Which moons show evidence of cryovolcanism?
18. Discuss evidence supporting the idea that Europa might have a subsurface ocean of liquid water.
19. Titan contains abundant amounts of methane. What process destroys methane in this moon's atmosphere?
20. In certain ways, Titan resembles a frigid version of the early Earth. Explain the similarities.
21. Some moons display signs of geological activity in the past. Identify some of the evidence for past activity.
22. Why do the outer planets but not the inner planets have rings? Describe a ground-based technique that led to the discovery of rings around the outer planets.
23. What are ring arcs and where are they found?
24. Identify and explain two possible mechanisms that can produce planetary ring material.
25. Explain two mechanisms that create gaps in Saturn's bright-ring system.
26. Describe ways in which diffuse rings differ from other planetary rings.
27. In Chapter 1, we stated that "all scientific theories are provisional." Explain how the discovery of the detailed structure of Saturn's F Ring (as described in the Process of Science Figure) challenged a scientific theory, and how this apparent conflict was ultimately resolved.
28. Astronomers think that most planetary rings eventually dissipate. Explain why the rings do not last forever. Describe and explain a mechanism that keeps planetary rings from dissipating.
29. Name one ring that might continue to exist indefinitely, and explain why it could survive when others might not.
30. Make a case for sending a space mission to one of the moons. Which moon would you choose to explore, and what types of observations would you try to obtain?

Applying the Concepts

31. Io has a mass of 8.9×10^{22} kg and a radius of 1,820 km.
    a. Using the formula provided in Working It Out 11.3, calculate Io's escape velocity.
    b. How does Io's escape velocity compare with the vent velocities of 1 km/s from its volcanoes?
32. Use the value of P^2/A^3 for Europa, as in Working It Out 11.1, to compute the mass of Jupiter.
33. Follow Working It Out 11.1 to compute the mass of Saturn using one of its moons.
34. Study Figure 11.2.
    a. Are the scales on (a) and (b) linear or logarithmic?
    b. About how much larger is the space shown in (b) than in (a)?
35. Planetary scientists have estimated that Io's extensive volcanism could be covering the moon's surface with lava and ash to an average depth of up to 3 millimeters (mm) per year.
    a. Io's radius is 1,820 km. If you assume Io is a sphere, what are its surface area and volume?
    b. What is the volume of volcanic material deposited on Io's surface each year?
    c. How many years would it take for volcanism to perform the equivalent of depositing Io's entire volume on its surface?
    d. How many times might Io have "turned inside out" over the age of the Solar System?
36. Consider the formula for tidal forces in Working It Out 11.2. If the radius of the moon increases but its mass stays the same, what happens to the tidal force? If the radius of the moon's orbit decreases, what happens to the tidal force? If the mass of the central planet increases, what happens to the tidal force?
37. Follow Working It Out 11.2 to compare the tidal force between Jupiter and Io with the tidal force between Earth and its Moon.
38. Imagine that a 60-kg astronaut is spacewalking outside the International Space Station, 380 km above Earth. Follow Working It Out 11.2 to find the tidal force on the astronaut, assuming she is oriented with her feet toward the center of Earth.
39. Assuming that all other numbers are held constant, make a graph of the tidal force versus the distance between a planet and its moon. On the same graph, plot the gravitational force (which falls off like $1/d^2$). Compare the two graphs to determine the relative importance of tidal and gravitational forces at various distances.
40. Particles at the very outer edge of Saturn's A Ring are in a 7:6 orbital resonance with the moon Janus. If the orbital period of Janus is 16 hours 41 minutes ($16^h\ 41^m$), what is the orbital period of the outer edge of Ring A?
41. Follow Working It Out 11.3 to find the escape velocity from Saturn's moon Janus.
42. The inner and outer diameters of Saturn's B Ring are 184,000 and 235,000 km, respectively. If the average thickness of the ring is 10 meters and the average density is 150 kilograms per cubic meter (kg/m^3), what is the mass of Saturn's B Ring?

43. The mass of Saturn's small, icy moon Mimas is 3.8×10^{19} kg. How does this mass compare with the mass of Saturn's B Ring, as calculated in question 42? Why is this comparison meaningful?

44. The inner and outer diameters of Saturn's B Ring are 184,000 and 235,000 km, respectively. Use this information to find the ratio of the periods of particles at these two diameters. Does the B Ring orbit like a solid disk or like a collection of separate particles?

45. Consider the escape velocity equation in Working It Out 11.3. For more massive planets, is the escape velocity higher or lower? For larger planets, is the escape velocity higher or lower? If you know the escape velocity of a planet, what other piece of information do you need in order to find the planet's mass?

USING THE WEB

46. Go to *Sky & Telescope*'s "Jupiter's Moons" Web page (http://www.skyandtelescope.com/wp-content/observing-tools/jupiter_moons/jupiter.html). Enter your date and time. Where are the four Galilean moons? Keep clicking on "[plus]1 hour" to see how their positions change over time. Which moon passes in front of (transits) Jupiter? If possible, observe these moons for a couple of nights through a small telescope, binoculars, or telephoto camera lens. Sketch the positions of the moons.

47. Look at the updated lists of giant planet moons on NASA's "Our Solar System" website (http://solarsystem.nasa.gov/planets; click on the planet's name and "Moons") or on the Carnegie Institution of Washington Department of Terrestrial Magnetism's (DTM) "Jupiter Satellite Page" (http://www.dtm.ciw.edu/users/sheppard/satellites). What are two of the more recently discovered moons on one of the planets? Are the orbits retrograde? What are the eccentricities and inclinations of the orbits? Where would these new moons fit on the graph of orbits on the DTM's website? Where did the names come from? Why are some of the more recent moons labeled "provisional"?

48. Go to the website for the *Cassini* mission (http://saturn.jpl.nasa.gov). Is *Cassini* still making observations? Look at a *Cassini* image of one of Saturn's moons. What does this image reveal about Saturn? Watch the video at http://saturn.jpl.nasa.gov/video/videodetails/?videoID=232 to listen to the "hiss" from the aurora. What moon is the cause of Saturn's aurora?

49. Missions to the moons:
   a. Do a search for the European Space Agency *JUICE* (*JUpiter ICy moons Explorer*) mission, scheduled for launch in the mid-2020s. Which moons will this mission study? What are the goals of the mission? What is the status of this project?
   b. Go to the "Destination Europa" website (http://europa.seti.org/). What are they advocating for? Do a search on "Europa Clipper" and "Europa mission" to see the status of proposed NASA space missions to Europa. Have any of them been approved for funding? What will the mission study?

50. Do a search to see if moons and rings have been confirmed on extrasolar planets. Why is this of interest to astronomers?

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

digital.wwnorton.com/astro5

Part A: Finding the Image Scale

Finding the scale of an image is like finding the scale on a map. On a map, each inch or centimeter represents miles or kilometers of actual space. The same thing is true in an image. If you take a picture of a meter stick and then measure the meter stick in the picture to be 10 cm long, you know that 10 cm in the picture represents 1 meter of actual space, and 1 cm in the picture represents 10 cm of actual space.

To find the scale, you must compare the size of something in the image with its actual size in space. In **Figure 11.35**, the moon Io is the known object.

1 Use a ruler to measure the diameter of this image of Io in millimeters.

---

2 Estimate the error in your measurement. (How far off could your measurement be?)

---

3 Find the radius from this diameter.

---

4 Look up the actual radius of Io (in kilometers) in Appendix 4 or online.

---

5 Find the image scale, s, as follows:

$$s = \frac{\text{Actual size of object}}{\text{Size of object on the image}}$$

---

6 What are the units of this image scale?

Part B: Finding the Sizes of Features on Io

7 There is a geyser near the center of Io, surrounded by a black circle and a white ring. This is Prometheus Patera. What is the diameter of the black ring around Prometheus in this image (in millimeters)? (Do not use the inset image.)

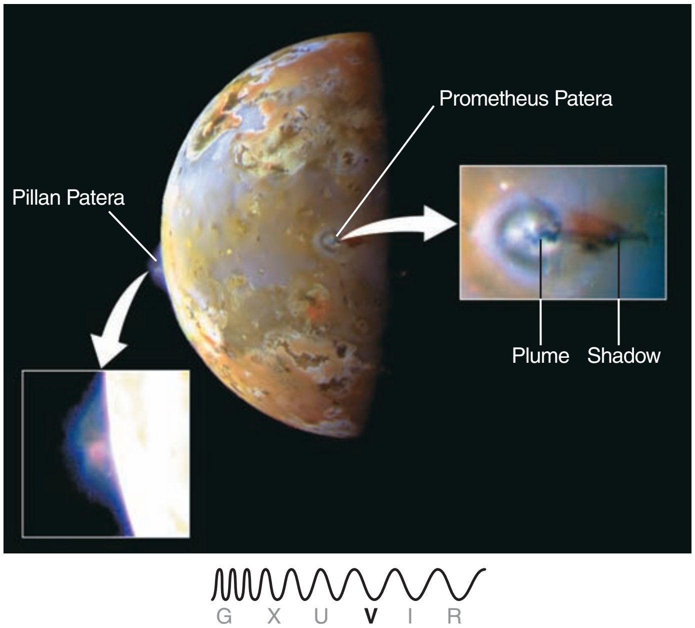


Figure 11.35 A *Galileo* image of Jupiter's moon Io, a very active moon. This image offers an opportunity to make some measurements.

8 Multiply the measured diameter by the image scale to find the actual size of the circle around Prometheus. Identify something on Earth that is about the same size.

---

9 On the limb of the moon, there is a purple plume from the eruption of a sulfur geyser. Follow steps 7 and 8 to find the height of that plume. Identify something on Earth that is about the same height.

There are at least two sources of error in this measurement. One is the error of measurement—how well you use a ruler, and how accurately you determined the top and the bottom of the plume. The other source of error is uncertainty about whether the plume is *exactly* at the limb of the moon. If the eruption occurred on the far side of Io, would your calculation be an overestimate or underestimate of the plume height?

12 Dwarf Planets and Small Solar System Bodies

In this chapter, we explore the small bodies remaining from the formation of the Solar System (see Chapter 7). These bodies formed today's dwarf planets, irregular moons, asteroids, and comets. These remaining planetesimals, and the fragments that some of them continually create, have revealed much about the physical and chemical conditions of the earliest moments in the history of the Solar System and how the Solar System formed and evolved. In addition, these planetesimals are important because some fraction of the water, gases, and organic material found on Earth and in the inner Solar System came from comets and asteroids.

LEARNING GOALS

By the conclusion of this chapter, you should be able to:

LG 1 List the categories of small bodies and identify their locations in the Solar System.

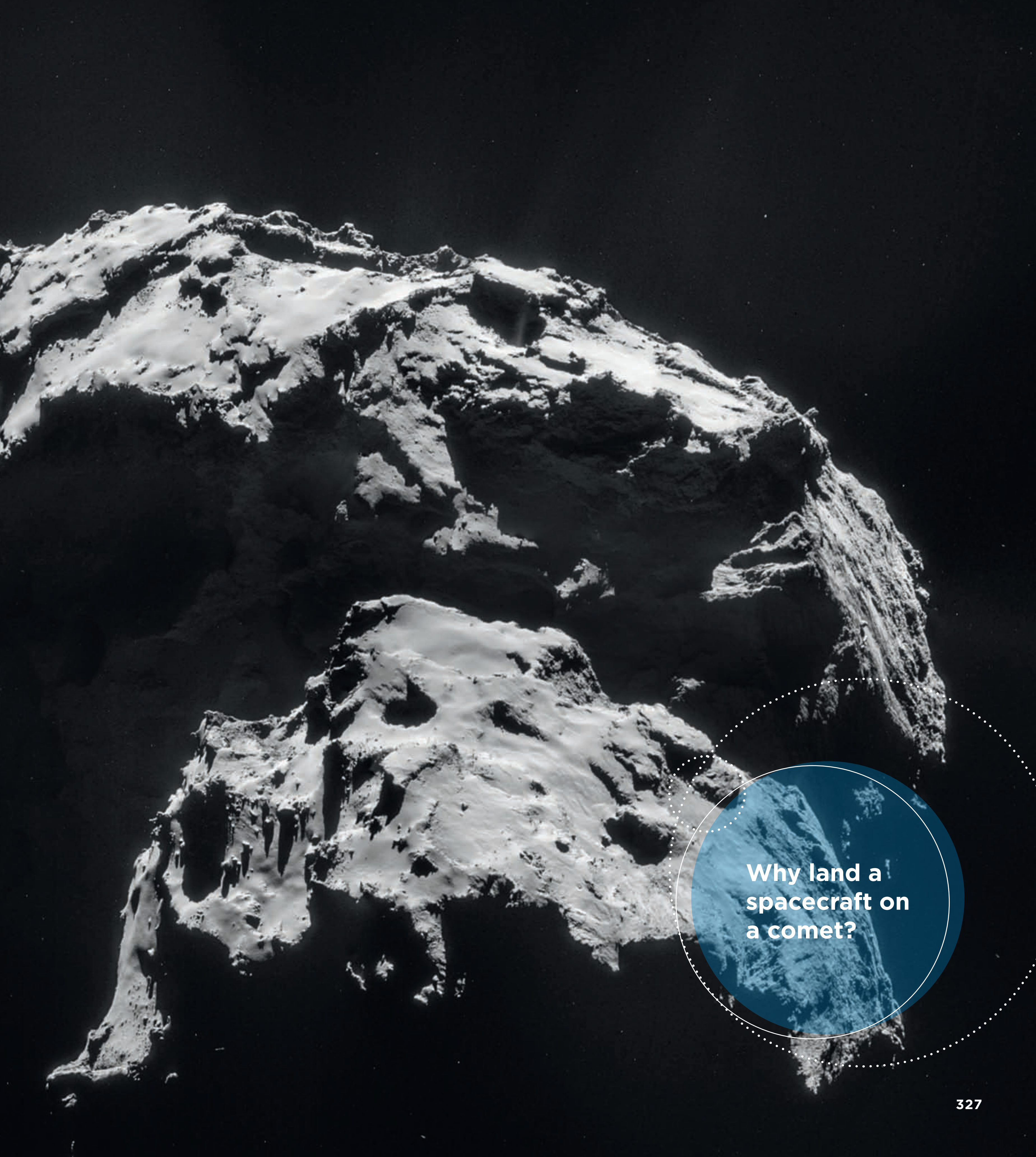
LG 2 Describe the defining characteristics of the dwarf planets in the Solar System.

LG 3 Describe the origin of the different types of asteroids, comets, and meteorites.

LG 4 Explain how asteroids, comets, and meteoroids provide important clues about the history and formation of the Solar System.

LG 5 Describe what has been learned from observations of recent impacts in the Solar System.

The European Space Agency landed a probe on Comet 67P/Churyumov-Gerasimenko in 2014. ▶▶▶


Why land a spacecraft on a comet?

12.1 Dwarf Planets May Outnumber Planets

Recall from Chapter 7 that very early in the history of the Solar System—at the same time that the Sun was becoming a star—tiny grains of primitive material stuck together to produce swarms of small bodies called *planetesimals*. Those that formed in the hotter, inner part of the Solar System were composed mostly of rock and metal; those in the colder, outer part were composed of ice, organic compounds, and rock. Some of the objects collided to become planets and moons. However, many are still present, and they remain a small but scientifically important component of the present-day Solar System.

Dwarf planets, asteroids, Kuiper Belt objects, comets, and meteoroids are smaller than planets and orbit the Sun. Dwarf planets are found in the asteroid belt and in the **Kuiper Belt**. The asteroid belt in the region between the orbits of Mars and Jupiter contains most of the asteroids in the Solar System. The Kuiper Belt is a disk-shaped population of comet nuclei extending from Neptune's orbit to perhaps thousands of astronomical units from the Sun.

The dwarf planets orbit the Sun and have round shapes, but because they have relatively small mass, they have not cleared the area around their orbits. As of this writing, there are five officially recognized dwarf planets in the Solar System: Pluto, Eris, Haumea, Makemake, and Ceres (their properties are tabulated in Appendix 4). Ceres is a large object in the main asteroid belt; the other dwarf planets are found in the Kuiper Belt. There are many dwarf planet candidates, but their shapes have not yet been measured well enough for certain classification.

Pluto

Throughout the 19th century, discrepancies were observed between the observed and predicted orbital positions of Uranus and Neptune. Early in the 20th century, astronomers hypothesized that an unseen body was perturbing the orbits of these planets. They called this body Planet X and estimated that it had 6 times Earth's mass and was located beyond Neptune's orbit. Astronomer Clyde W. Tombaugh (1906–1997) discovered Planet X in 1930, not far from its predicted position. It became the Solar System's ninth planet and was named Pluto for the Roman god of the underworld. However, observational evidence soon indicated that the mass of Pluto was far too small to have produced the perturbations in the orbits of Uranus and Neptune. When astronomers reanalyzed the 19th century observations, they found that the orbital "discrepancies" were a mistake. Pluto's discovery thus turned out to be a coincidence.

Pluto's orbit is 248 Earth years long and quite elliptical, and it is tilted with respect to the plane of the Solar System. Its orbit periodically crosses inside of Neptune's nearly circular orbit—from 1979 to 1999, Pluto was closer to the Sun than Neptune. Pluto is only two-thirds as large as our Moon. Pluto has five known moons, the largest of which is Charon, about half the size of Pluto. The total mass of the Pluto-Charon system is 1/400 that of Earth, or 1/5 the mass of the Moon. As with Uranus, the plane of Pluto's equator is nearly perpendicular to its orbital plane. Pluto and Charon are a tidally locked pair: each has one hemisphere that always faces the other (**Figure 12.1a**).

Pluto and Charon were not on *Voyager*'s route through the Solar System, so until recently there was limited information about their surface properties or geological history. This changed when NASA's *New Horizons* flyby spacecraft passed within 12,500 km to Pluto in July 2015. Astronomers were surprised by the varied surface features seen in *New Horizons* images of Pluto, including a large bright region primarily composed of carbon monoxide ice (Figure 12.1b). Figure 12.1c shows 3500-meter high ice mountains and icy plains. Pluto's surface contains an icy mixture of frozen water, carbon dioxide, methane, and carbon monoxide, with flowing nitrogen ice. Pluto has a thin atmosphere of nitrogen, methane, ethane, and carbon monoxide: these gases freeze out of the atmosphere when Pluto is more distant from the Sun and therefore cooler. Charon has no atmosphere. Its surface has deep canyons, which might have formed as an ancient ocean froze and pushed surface outwards (Figure 12.1d). Few craters are seen on either body, which suggests the surfaces are relatively young, but it is not yet known if there was recent geological activity or impacts.

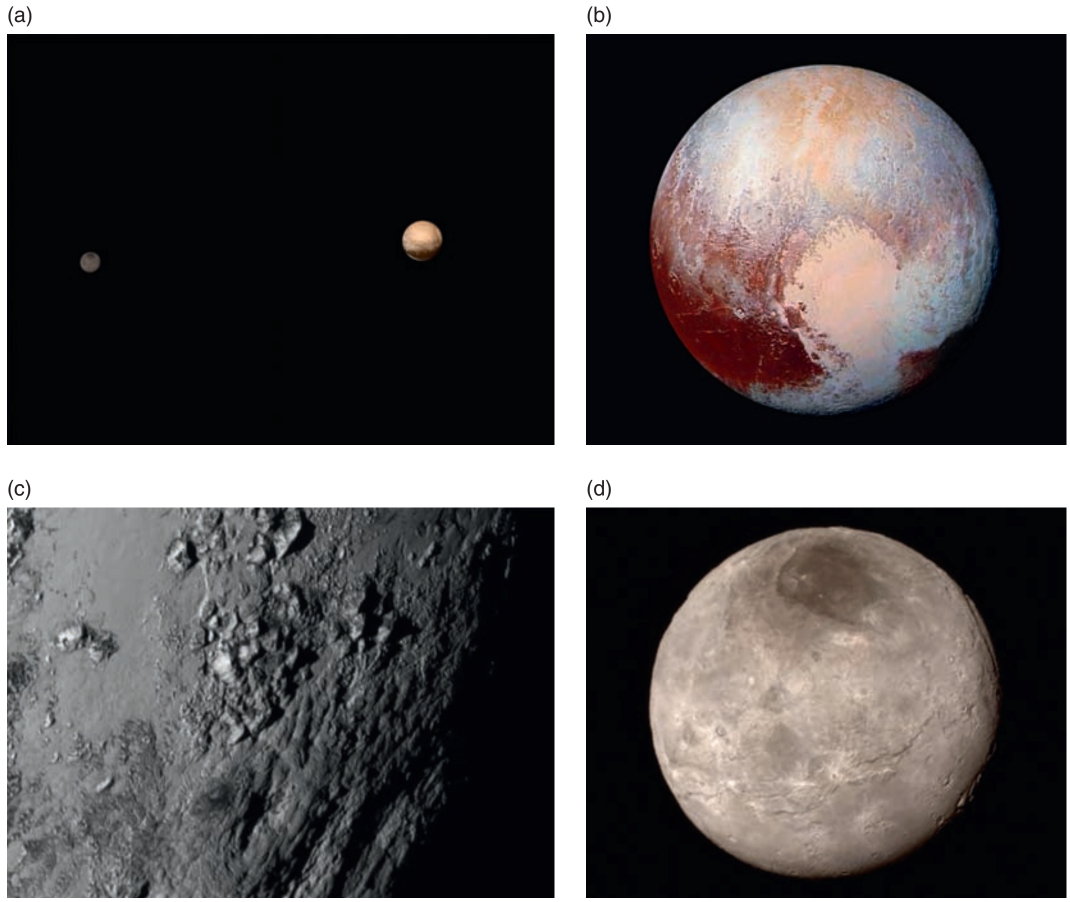


Figure 12.1 (a) Pluto (right) and Charon (to scale). All images are from the New Horizons spacecraft flyby in 2015. (b) Enhanced color image of Pluto. (c) Young mountains of ice on Pluto rise to 3500m, suggesting recent geological activity. (d) Image of Pluto's largest moon Charon.

As astronomers discovered more about Pluto and other objects beyond Neptune's orbit, some questioned Pluto's classification as a planet, and a debate ensued. In 2005, astronomers identified an object more distant than Pluto, later named Eris, and then Eris's moon, Dysnomia. Observations of Dysnomia's orbit yielded a mass for Eris, which turned out to be about 28 percent greater than Pluto's mass. Pluto and Eris have similar nitrogen and methane abundances and a relatively large moon. At this point, the inevitable question emerged: Should astronomers consider Eris to be the Solar System's tenth planet? Or should neither Pluto nor Eris be called a planet? The International Astronomical Union (IAU) made its decision in August 2006 (**Process of Science Figure**): Pluto is round like the classical planets, but it is not able to clear its neighborhood, so it was reclassified as a dwarf planet (details of the IAU resolution can be found in Appendix 9).

Eris, Haumea, and Makemake

These three dwarf planets were discovered in the 21st century. Eris is about the same size as Pluto but is more massive. Eris also has a relatively large moon, called Dysnomia. The highly eccentric orbit of Eris shown in **Figure 12.2** carries it from 37.8 astronomical units (AU) out to 97.6 AU away from the Sun, with an orbital

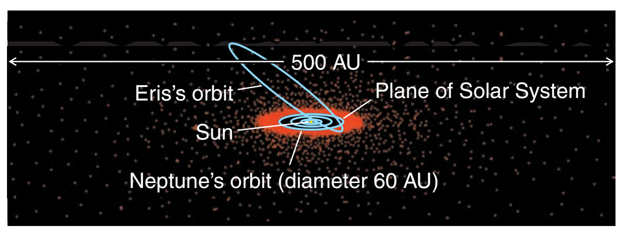
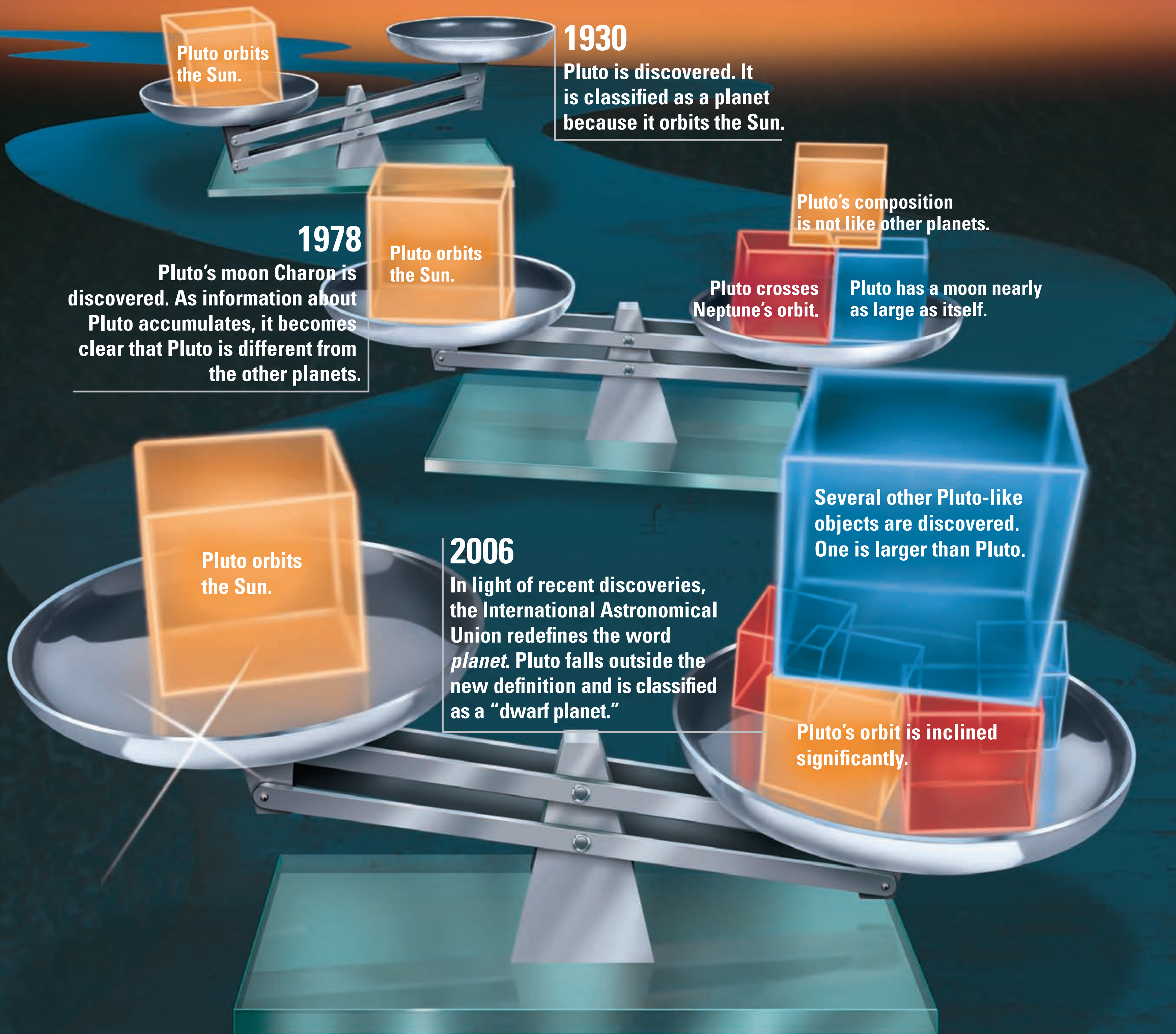


Figure 12.2 Eris's orbit is both highly eccentric and highly inclined to the rest of the Solar System.

Process of Science

HOW TO CLASSIFY PLUTO

Pluto's reclassification from planet to dwarf planet in 2006 got a lot of publicity. The reclassification of Pluto is a clear example of the scientific method in practice.



Scientific decision making must follow the weight of the evidence.

12.1 Working It Out Eccentric Orbits

Many of the objects discussed in this chapter are far from the Sun, and their complete orbits take many years. But observing a complete orbit is not necessary for determining an object's semimajor axis and eccentricity: these values can be obtained from watching how the object moves in just a fraction of its orbit. Astronomers can calculate the orbits of distant objects as they approach the Sun in a highly elliptical orbit and determine if they will come near to Earth.

Recall Kepler's law for objects orbiting the Sun: $P^2 = A^3$. **Figure 12.3** shows how eccentricity is defined mathematically as the distance from the center of the orbit to one focus (the Sun) divided by the semimajor axis (A). (The orbits of the planets in the Solar System range from nearly circular for Venus to an eccentricity [e] of 0.2 for Mercury.) The types of objects discussed in this chapter generally have higher eccentricities.

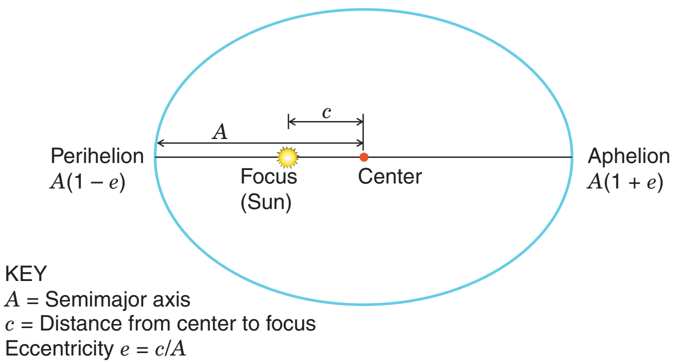


Figure 12.3 This drawing of an elliptical orbit shows eccentricity, aphelion, and perihelion.

We can relate the eccentricity to the closest approach and the farthest distance in the orbit, as seen in Figure 12.3, so that the object's closest approach to the Sun, its **perihelion**, equals $A(1 - e)$, and the object's farthest distance from the Sun, its **aphelion**, equals $A(1 + e)$. So if we know the semimajor axis and eccentricity of an orbit, we can calculate how close to and how far away from the Sun an object's orbit takes it.

For a first example, consider the orbit of dwarf planet Eris. The eccentricity of Eris's orbit is 0.44, and the semimajor axis of its orbit (A) is 67.7 AU. Therefore, we can calculate how close to the Sun and how far away Eris gets as follows:

$$\text{Perihelion} = A(1 - e) = 67.7(1 - 0.44) = 67.7 \times 0.56 = 37.9 \text{ AU}$$

$$\text{Aphelion} = A(1 + e) = 67.7(1 + 0.44) = 67.7 \times 1.44 = 97.5 \text{ AU}$$

At this time, Eris is close to aphelion. But when it approaches perihelion, it will cross the orbit of Pluto, whose distance varies from 29.7 to 48.9 AU.

For a second example, let's look at Apollo asteroid 2005 YU55, which has a semimajor axis of 1.14 AU and an orbital eccentricity of 0.43. We can calculate its perihelion and aphelion similarly:

$$\text{Perihelion} = A(1 - e) = 1.14(1 - 0.43) = 1.14 \times 0.57 = 0.65 \text{ AU}$$

$$\text{Aphelion} = A(1 + e) = 1.14(1 + 0.43) = 1.14 \times 1.43 = 1.63 \text{ AU}$$

These results indicate that the orbit of 2005 YU55 crosses the orbits of Earth and Mars. In November 2011, this asteroid passed 324,900 kilometers (km) from Earth—which is about 85 percent of the distance to the Moon.

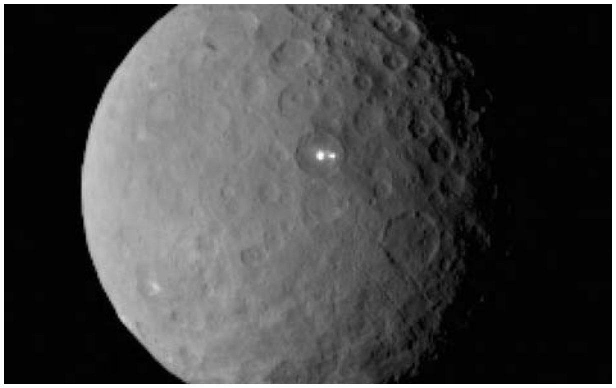
period of 557 years (**Working It Out 12.1**). Eris is near the most distant point in its orbit, making it the most remote known object in the Solar System. At this distance it is about 100 times fainter than Pluto. The eccentric orbits of other Solar System bodies will eventually carry them farther away than Eris, however, so Eris will not always be the most distant object known.

When astronomers combine the observed brightness of Eris with its diameter, they find it has a surprisingly high albedo of 0.96. (Recall from Chapter 5 that albedo is a measure of how much light an object reflects.) The surface of Eris is more highly reflecting than that of any other major Solar System body except Enceladus, so Eris too must have a coating of pristine ice. The surface of Enceladus is water ice, while Eris is covered with methane ice. At its present location, the average surface temperature on Eris is cold enough to freeze out any atmospheric methane, but it will probably develop a methane atmosphere when it comes closest to the Sun in the year 2257.

Haumea and Makemake are both smaller and have slightly larger orbits than that of Pluto (**Figure 12.4**). Haumea has two

Figure 12.4 This NASA illustration shows the five dwarf planets compared to our Moon (Luna) and Earth.

(a)



(b)

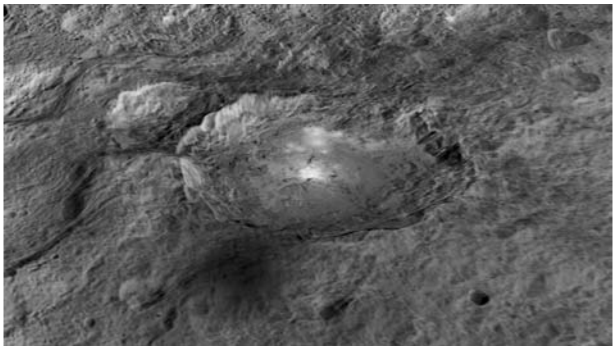


Figure 12.5 (a) Dwarf planet Ceres photographed by the *Dawn* mission. (b) Haze has been observed above the white spots, which suggests they are made of ice.

moons—Hi'iaka and Namaka—enabling astronomers to calculate the system's mass. Although Haumea has sufficient mass to pull itself into a spherical shape, it spins so rapidly on its axis that its shape is flattened, with an equatorial radius that is approximately twice its polar radius. This difference between the equatorial and polar radii gives Haumea an oblateness (a measure of how far an object is from perfectly round) of 0.5, the most distorted shape of any of the planets or dwarf planets. HST infrared imaging indicates that Haumea and its two moons are covered in water ice. Astronomers think that these three objects and some smaller debris were left over after a larger body broke up following a collision. No moons have been discovered orbiting Makemake, so less is known about this dwarf planet than about Pluto, Haumea, or Eris.

Ceres

The dwarf planet Ceres was discovered in 1801 when Sicilian astronomer Giuseppe Piazzi found a bright object between the orbits of Mars and Jupiter. Piazzi named the new object Ceres. When Piazzi discovered Ceres, he thought he might have found a hypothetical "missing planet." But as more objects were discovered orbiting the region between Mars and Jupiter, astronomers classified Ceres as belonging to a new category of Solar System objects called asteroids. Ceres is the largest body in the main asteroid belt. It also is now called a dwarf planet, because although it is round (**Figure 12.5a**), it has not cleared its surroundings.

With a diameter of about 940 km, Ceres is larger than most moons but smaller than any planet. It contains about a third of the total mass in the asteroid belt, but only about 1.3 percent of the mass of Earth's Moon. Ceres rotates on its axis with a period of about 9 hours, typical of many asteroids. As a large planetesimal, Ceres seems to have survived largely intact, although it appears to have undergone differentiation at some point in its early history.

About a quarter of its mass exists in the form of a water-ice mantle that surrounds a rocky inner core. Water vapor coming from two locations on Ceres indicates that there is water ice in specific locations on the surface. NASA's *Dawn* mission went into orbit around Ceres in 2015. Observed geological features include a 5-km high mountain and craters 4-5 km deep. One crater has bright spots and haze within the boundaries of its rim, suggesting the spots are made of ice (Figure 12.5b). Dawn will remain in Ceres' orbit for the remainder of its mission.

CHECK YOUR UNDERSTANDING 12.1

Why are these objects called dwarf *planets* even though they are smaller than some moons?

12.2 Asteroids Are Pieces of the Past

After Piazzi found Ceres in 1801, a number of similar objects were discovered in the region between the orbits of Mars and Jupiter. Because these new objects appeared in astronomers' eyepieces as nothing more than faint points of light, William and Caroline Herschel (the brother-sister pair of astronomers who discovered Uranus) named them *asteroids*, a Greek word meaning "starlike." As the years went by, more asteroids were discovered, and there are now estimated to be 1 million to 2 million asteroids larger than 1 km in size, and many more that are smaller. Because these objects are in the Solar System, many of them move quickly enough across the sky that their motion is noticeable over a few hours. Both professional and amateur astronomers have discovered asteroids.

Recall that an asteroid is a primitive planetesimal that did not become part of the accretion process that formed planets. The planetesimals that formed our Solar System's planets and moons have been so severely modified by planetary processes that nearly all information about their original physical condition and chemical composition has been lost. By contrast, asteroids and comet nuclei constitute an ancient and far more pristine record of what the early Solar System was like. Asteroids are composed of the same type of rocky material that became the inner planets, and comets are composed of the same type of icy material that became the outer planets. Thus, planetary scientists study asteroids in order to learn about the inner planets and their formation. In this section, we will study the orbits and composition of the asteroids.

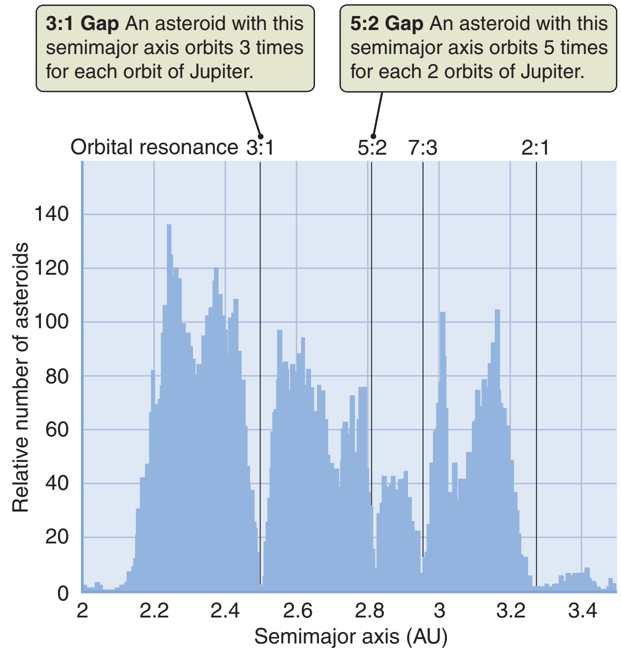


Figure 12.6 This plot shows the relative number of asteroids in the main belt with a given orbital semimajor axis. The gaps in the distribution of asteroids, called *Kirkwood gaps*, are caused by orbital resonances with Jupiter.

The Distribution of Asteroids

Asteroids are found throughout the Solar System. Most orbit the Sun in several distinct zones, with the majority residing between the orbits of Mars and Jupiter in the **main asteroid belt**. The main belt contains at least 1,000 objects larger than 30 km in diameter, of which about 200 are larger than 100 km. Although there are a great number of asteroids, they account for only a tiny fraction of the matter in the Solar System. Some of the asteroids are bound to another asteroid in a double system, and more than 200 asteroids have moons, some similar in size to the asteroids themselves. At least one asteroid has a ring.

Asteroids are not distributed randomly throughout the main asteroid belt: there are several empty regions. **Figure 12.6** shows that there are very few asteroids that orbit at specific distances from the Sun. These "gaps" in the asteroid belt are called Kirkwood gaps after Daniel Kirkwood (1814–1895), the astronomer who first recognized them. Recall the idea of orbital resonances discussed in the past chapter: the orbital periods of some moons around their planets are numerically related. Similarly, all of the Kirkwood gaps in the asteroid belt correspond to resonances: asteroid orbits that are related to the orbital period of Jupiter by the ratio of two small integers. The boundaries of the asteroid belt are set by some of these resonances. The inner boundary of the asteroid belt, at 1.8 AU, corresponds to the 5:1 orbital resonance of Jupiter; the outer boundary, at 3.3 AU, corresponds to the 2:1 orbital resonance.

To understand the Kirkwood gaps, consider the example of an asteroid starting with an orbital period exactly half that of Jupiter, a 2:1 orbital resonance. After two complete asteroid orbits, the asteroid, Jupiter, and the Sun are lined up in the same place where they started. As Jupiter and the asteroid continue in their orbits, they line up at this same location again and again, every 11.86 years (the orbital period of Jupiter). The gravitational force of the Sun on the asteroid is more than 360 times stronger than the gravitational force that Jupiter exerts on this asteroid at its closest approach. A single close pass between Jupiter and the asteroid does very little to the asteroid's orbit. For an asteroid that is *not* in orbital resonance with Jupiter, the tiny gravitational tugs from Jupiter come at a different place in its orbit each time. The effects of these random tugs average out, and as a result even multiple passes close to Jupiter have little overall effect. For an asteroid that has a 2:1 orbital resonance with Jupiter as in this example, the tug from Jupiter comes at the same place in its orbit every time.

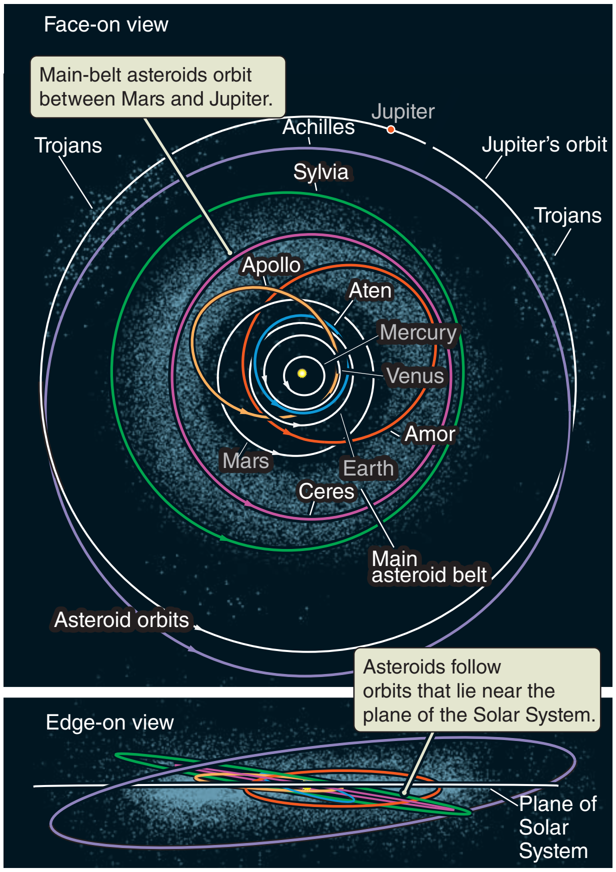


Figure 12.7 This illustration shows face-on and edge-on views of asteroid orbits. Blue dots show the locations of known asteroids at a single point in time. The orbits of Aten, Amor, and Apollo (prototype members of some groups of asteroids) are shown. Most asteroids, such as Sylvia, are main-belt asteroids. Achilles was the first Trojan asteroid to be discovered.

The repeated tugs from Jupiter at the same place add up, changing the asteroid's orbit. Thus, an asteroid in such an orbit would not stay there long. This is why there are no asteroids with orbital periods equal to half the orbital period of Jupiter. Other orbital resonances, such as a 3:1 resonance, will have a similar effect. The reason asteroids are not found in the Kirkwood gaps is that their gravitational interaction with Jupiter prevents them from staying there.

There are several groups of asteroids not in the main asteroid belt. As shown in **Figure 12.7**, these are divided according to their orbital characteristics. Trojans share Jupiter's orbit and are held in place by interactions with Jupiter's gravitational field. Three other groups are defined by their relationship to the orbits of Earth and Mars: Apollo asteroids cross the orbits of Earth and Mars, Aten asteroids cross Earth's orbit but not that of Mars, and Amor asteroids cross the orbit of Mars but not that of Earth. All three of these groups are named for a prototype that is representative of the group.

Asteroids whose orbits bring them within 1.3 AU of the Sun are called **near-Earth asteroids**. These asteroids, along with a few comet nuclei, are known collectively as **near-Earth objects (NEOs)**. NEOs occasionally collide with Earth or the Moon. Astronomers estimate that between 500 and 1,000 NEOs have diameters larger than a kilometer. Collisions with such objects are geologically important and have dramatically altered the history of Earth and life on Earth as discussed in Chapter 8.

Part of NASA's mission is to identify and track NEOs. NASA's Wide-field Infrared Survey Explorer (WISE), an infrared telescope in space, surveyed the entire sky during 2010. The data suggest there are about 20,000 mid-sized asteroids (100 meters to 1 km in diameter) near Earth. WISE also observed more than 150,000 asteroids in the main belt, including 33,000 new ones, as well as 2,000 Jovian Trojans. The WISE mission was reactivated in late 2013 and is currently searching for NEOs.

The Composition and Classification of Asteroids

Most asteroids are relics of rocky or metallic planetesimals that originated in the region between the orbits of Mars and Jupiter. Although early collisions between these planetesimals created several bodies large enough to differentiate, Jupiter's tidal disruption and possible orbital migration prevented them from forming a single Moon-sized planet. As they orbit the Sun, asteroids continue to collide with one another, producing small fragments of rock and metal. Most meteorites are pieces of these asteroidal fragments that have found their way to Earth and crashed to its surface.

With a few exceptions, the mass of an asteroid is too small for self-gravity to have pulled it into a spherical shape. Some asteroids have highly elongated irregular shapes, somewhat like potatoes, suggesting objects that either are fragments of larger bodies or were created haphazardly from collisions between smaller bodies. Astronomers have measured the masses of a number of asteroids by noting the effect of their gravity on Mars, on spacecraft passing nearby, on each other, or by the orbits of their moons. The total mass of the asteroids in the main belt is estimated to be about 3 times the mass of Ceres, or 4 percent the mass of the Moon. Their densities can be found from their mass and size and range between 1.3 and 3.5 times the density of water. The lower-density asteroids are shattered heaps of rubble, with large voids between the fragments.

Asteroids rotate just as planets and moons do, although irregularly shaped asteroids can wobble a lot as they spin. The rotation periods of some asteroids range from 2 hours to longer than 40 Earth days. Rotation periods for asteroids are measured by watching changes in their brightness as they alternately present their broad and narrow faces to Earth. Different groups of asteroids have different average rotations.

Asteroids can also be classified by composition. Meteorites found on Earth come from asteroids, which come from planetesimals, as shown in **Figure 12.8**. As larger planetesimals accreted smaller objects, thermal energy from impacts and the decay of radioactive elements heated them. Despite this heating, some planetesimals never reached the high temperatures needed to melt their interiors: they simply cooled off. They look like rubble piles, pretty much as they were when they formed. These planetesimals, the most common type of asteroid in the main belt, are called **C-type** (carbon) asteroids. They are composed of primitive material that has largely been unmodified since the origin of the Solar System almost 4.6 billion years ago.

In contrast, some planetesimals were heated enough by impacts and radioactive decay to cause them to melt and differentiate, with denser matter such as iron sinking to their centers. Lower-density material—such as compounds of calcium, silicon, and oxygen—floated toward the surfaces of these planetesimals and combined to form mantles and crusts of silicate rocks. **S-type** (stony) asteroids may be pieces of the mantles and crusts of such differentiated planetesimals and are

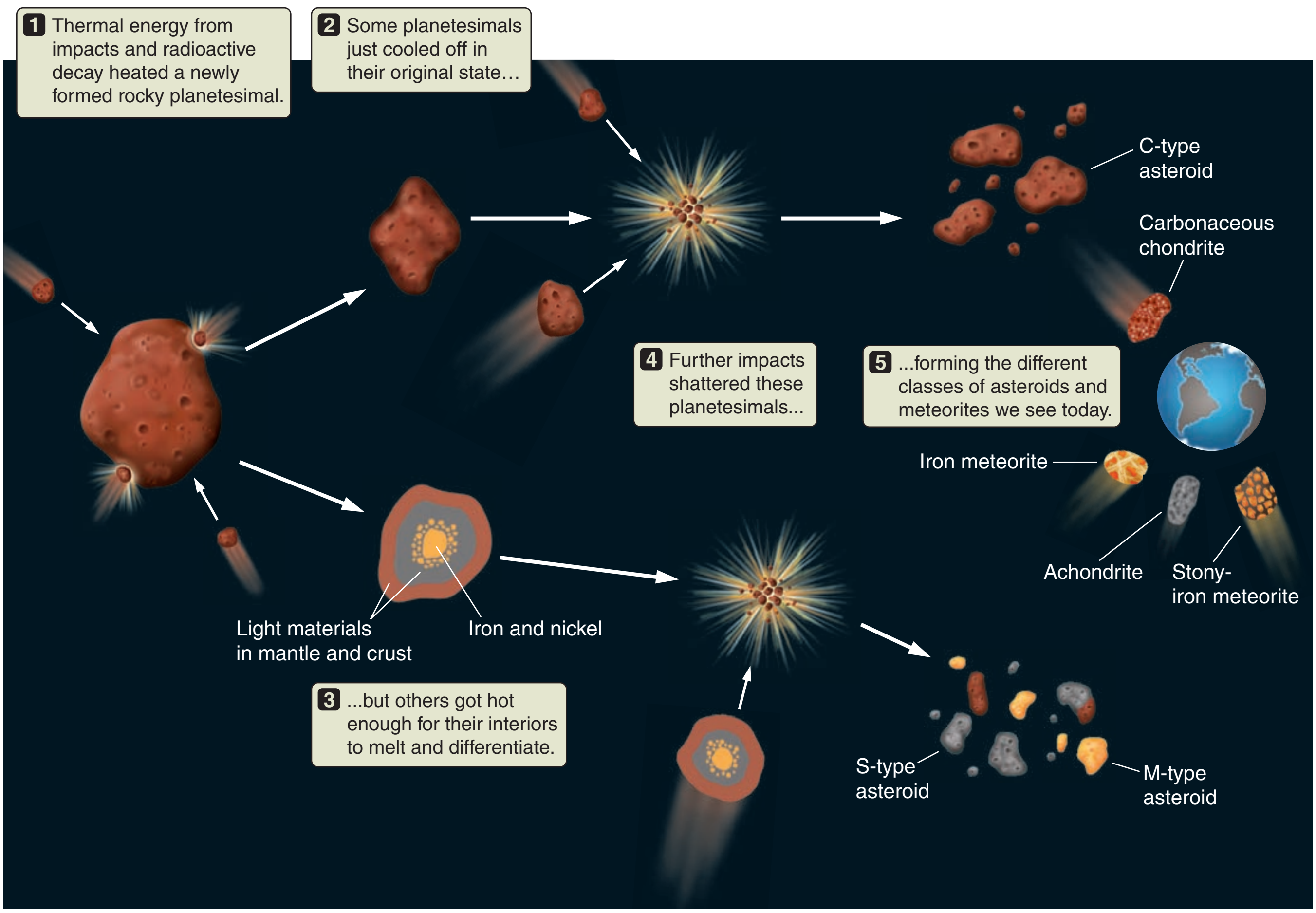


Figure 12.8 The fate of a rocky planetesimal in the young Solar System depends on whether it gets large and hot enough to melt and differentiate, as well as on the impacts it experiences. Different histories led to the different types of asteroids and meteorites found today. (Images not to scale.)

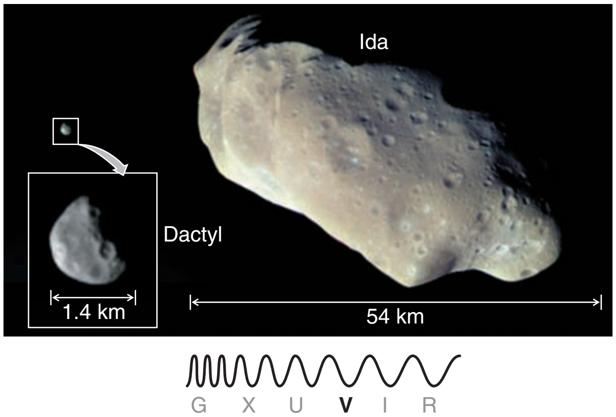


Figure 12.9 This *Galileo* spacecraft image shows the asteroid Ida with its tiny moon, Dactyl (shown enlarged in the inset).

chemically similar to volcanic rocks found on Earth. They were hot enough at some point to lose their carbon compounds and other volatile materials to space. Similarly, **M-type** (metal) asteroids are fragments of the iron- and nickel-rich cores of one or more differentiated planetesimals that shattered into small pieces during collisions with other planetesimals.

Recently, some asteroids have been shown to have ice on their surface. Using a ground-based infrared telescope, astronomers found ice on 24 Themis: it is one of the largest main-belt asteroids (diameter 200 km) and orbits the Sun at the outer edge of the asteroid belt. Water ice covers its entire surface, and organic molecules were also found there. Hydrated minerals have been found on meteorites thought to have come from outer-main-belt asteroids, but this was the first direct detection of water ice on an asteroid. This discovery may indicate that there is a continuum rather than a strict boundary between icy comets and rocky asteroids. The observations support the idea that both asteroids and comets brought water and organic material to the early Earth.

Asteroids Viewed Up Close

Several asteroids have been visited by spacecraft. In 1991, the *Galileo* mission passed by two S-type asteroids while on its way to Jupiter. The small asteroid Gaspra is cratered and irregular in shape, about 18 × 11 × 9 km in size. Faint, groovelike patterns may be fractures from the impact that chipped Gaspra from a larger planetesimal. Distinctive colors imply that Gaspra is covered with a variety of rock types. Later, *Galileo* passed close to asteroid Ida in the outer part of the main asteroid belt (**Figure 12.9**). *Galileo* flew so close to Ida that its cameras could see details as small as 10 meters across. Ida is 60 × 19 × 25 km in size, and its surface is about a billion years old, twice the age estimated for Gaspra. Like Gaspra, Ida contains fractures, indicating that these asteroids must be made of relatively solid rock. This supports the idea that some asteroids are chips from larger, solid objects. The *Galileo* images also revealed a tiny moon orbiting Ida, called Dactyl, which is only 1.4 km across and cratered from impacts.

The first spacecraft to land on an asteroid was *NEAR Shoemaker*, which was gently crash-landed into asteroid Eros in 2002 after a year of taking observations. Chemical analyses confirmed that the composition of Eros is like that of primitive meteorites. In November 2005, the Japanese spacecraft *Hayabusa* made contact with the small (less than 0.5 km) S-type asteroid Itokawa. *Hayabusa* collected small samples of dust that were returned to Earth in 2010—the first sample-return mission from an asteroid. Chemical analysis showed that such S-type asteroids are the parents of a type of meteorite found on Earth. They also suggested that Itokawa had been much larger, more than 20 km in diameter, when it formed.

Figure 12.10 This image of Vesta was taken by *Dawn* in 2012. Its north pole is in the middle of the image.

In 2011, NASA's *Dawn* spacecraft went into orbit around Vesta (**Figure 12.10**), the second most massive body in the asteroid belt (after Ceres). Vesta is small (525 km in diameter) compared to the terrestrial planets but large compared to the other visited asteroids. The data from *Dawn* indicate that Vesta is a leftover intact protoplanet that formed within the first 2 million years of the condensation of the first solid bodies in the Solar System. It has an iron core and is differentiated, so it is more like the planets than like other asteroids.

Vesta's spectrum matches the reflection spectrum of a peculiar group of meteorites that look like rocks taken from iron-rich lava flows on Earth and the Moon. A collision—or two—that created the two large impact basins in the south polar region of Vesta (**Figure 12.11a**) blasted material into space that then landed on Earth as these meteorites. These basins are only 1 billion to 2 billion years old. The younger basin is 500 km across and 19 km deep—a depth greater than the height of Mauna Kea in Hawaii. Smaller adjacent impact craters in the northern hemisphere (Figure 12.11b) were nicknamed "Snowman."

CHECK YOUR UNDERSTANDING 12.2

Remnants of volcanic activity on the asteroid Vesta indicate that members of the asteroid belt: (a) were once part of a single protoplanet that was shattered by collisions; (b) have all undergone significant chemical evolution since formation; (c) occasionally grow large enough to become differentiated and geologically active; (d) used to be volcanic moons orbiting other planets.

12.3 Comets Are Clumps of Ice

Early cultures viewed the sudden and unexpected appearance of a bright comet as an omen. Comets were often seen as dire warnings of disease, destruction, and death, but sometimes as portents of victory in battle or as heavenly messengers announcing the impending birth of a great leader. The earliest records of comets date from as long ago as the 23rd century BCE. Until the end of the Middle Ages, comets were regarded as mysterious temporary atmospheric phenomena rather than as astronomical objects. In the 16th century, Tycho Brahe reasoned that if comets were atmospheric phenomena like clouds, then their appearance and location in the sky should be very different to observers located many miles apart. But when Tycho compared sightings of comets made by observers at several different sites, he found no evidence of such differences, and he concluded that comets must be at least as far away as the Moon.

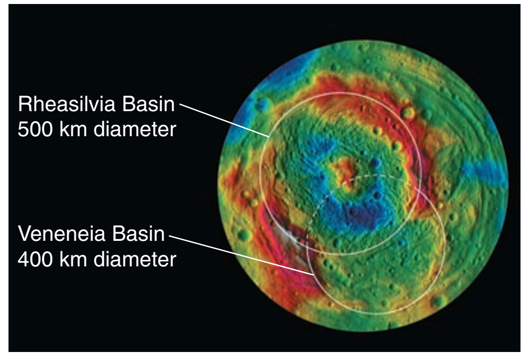
Today, we know that comets are icy planetesimals that formed from primordial material. They spend most of their time adrift in the frigid outer reaches of the Solar System. Comet nuclei put on a show only when their orbit brings them deep enough into the inner Solar System to undergo destructive heating from the Sun—they emit streams of dust and gas. In this section, we will examine the orbits and composition of the comets.

The Homes of the Comets

A *comet* is a complex object consisting of a small, solid, icy nucleus, an atmospheric halo, and a tail of dust and gas: a comet nucleus is the "heart" of the comet and contains most of the comet's mass. When very distant from the Sun, the comet is entirely nucleus—frozen throughout. As it approaches the Sun, the coma forms first, and then the tail forms. When they are near enough to the Sun to show the effects of solar heating, they are called **active comets**, or often simply *comets*. Most of these icy bodies are much too small and far away to be seen, so no one really knows how many there are. Estimates for our Solar System range as high as a trillion (10^{12}) comet nuclei—more than the number of stars in the Milky Way Galaxy—but astronomers have seen only several thousand.

We know where comets come from by observing their orbits as they pass through the inner Solar System. Comets fall into two distinct groups named for scientists Gerard Kuiper (1905–1973) and Jan Oort (1900–1992).

(a)



(b)

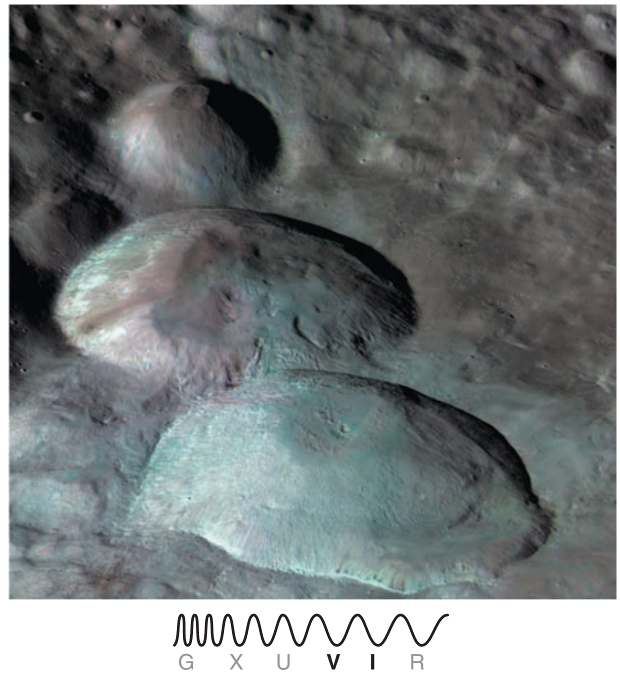


Figure 12.11 (a) This image shows impact basins at the south pole of Vesta. By counting the craters on top of it, astronomers estimate Rheasilvia to be 1 billion years old. Veneneia is partly beneath Rheasilvia and is estimated to be 2 billion years old. Red is higher elevation. (b) This image shows three craters in the northern hemisphere—60, 50, and 22 km across, respectively. The feature was nicknamed "Snowman."

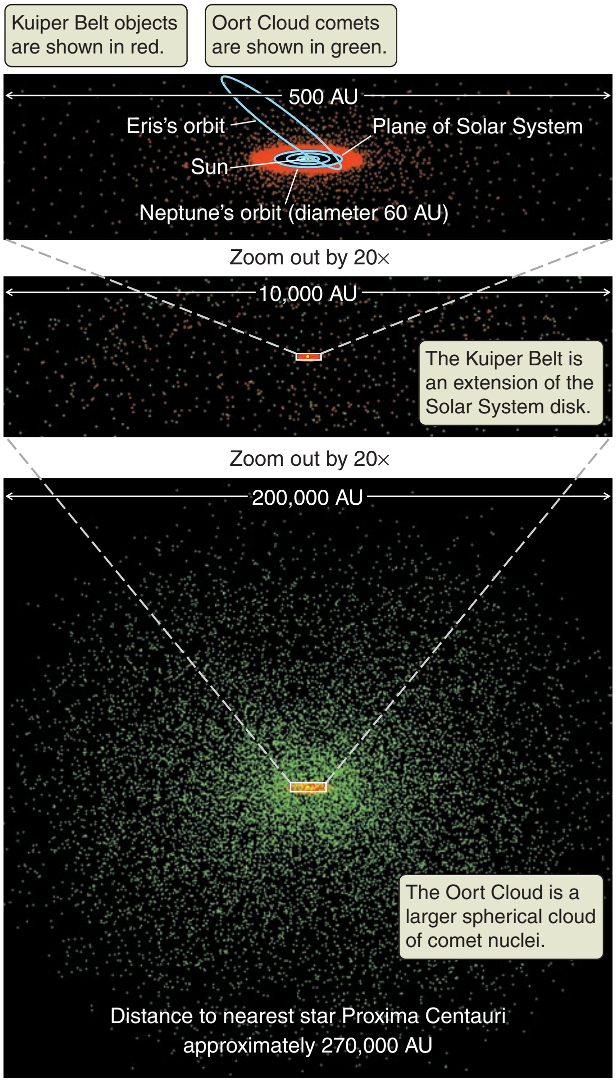


Figure 12.12 The top image shows that most comets near the inner Solar System populate an extension to the disk of the Solar System called the Kuiper Belt (red). The middle image zooms out to show the expanded Kuiper Belt. The bottom image zooms out to illustrate the spherical Oort Cloud, which is far larger and contains many more comet nuclei (green).

Kuiper Belt The Kuiper Belt is a disk-shaped population of comet nuclei that begins at about 30 AU from the Sun, near the orbit of Neptune, and extends outward to about 50 AU (**Figure 12.12**). Comets from the Kuiper Belt orbit the Sun in a disk-shaped region aligned with the Solar System. The innermost part of the Kuiper Belt contains tens of thousands of icy planetesimals known as Kuiper Belt objects (KBOs), or sometimes as *trans-Neptunian objects* (*TNOs*). The largest KBOs are similar in size to Pluto and Eris. With a few exceptions, the sizes of KBOs are difficult to determine, because although brightness and approximate distance are known, their albedos are uncertain. Reasonable limits for the albedos can set maximum and minimum values for their size. Like asteroids, some KBOs have moons, and at least one has three moons. We know very little of the chemical and physical properties of KBOs because of their great distance. After it encountered Pluto in 2015, the *New Horizons* spacecraft is scheduled to continue outward into the Kuiper Belt, where it will be maneuvered to fly close to one or more KBOs.

One of the larger known KBOs, Quaoar (pronounced "kwa-whar"), is also one of the few whose size astronomers have independently measured—about 900 km. From its apparent brightness, distance, and size, astronomers calculate Quaoar's albedo to be 0.20, making it more reflective than the nuclei of those comets that have entered the inner Solar System but far less reflective than Pluto. Quaoar's remote location and pristine condition have apparently allowed some volatile ices to survive on its surface, including crystalline water ice, methane, and ethane. Quaoar has a nearly circular orbit about the Sun and has a small moon, which enables astronomers to estimate its mass.

The icy planetesimals in the Kuiper Belt are packed closely enough to interact gravitationally from time to time. In such events, one object gains energy while the other loses it. The "winner" may gain enough energy to be sent into an orbit that reaches far beyond the boundary of the Kuiper Belt. The "loser" may fall inward toward the Sun.

Oort Cloud Unlike the flat disk of the Kuiper Belt, the **Oort Cloud** is a spherical distribution of planetesimals (Figure 12.12) that are much too distant to be seen by even the most powerful telescopes. Astronomers determine the size and shape of the Oort Cloud from the orbits of this region's comets, which approach the Sun from all directions and from as far as 100,000 AU away—nearly halfway to the nearest stars.

Sedna is an interesting object in the inner Oort Cloud whose highly elliptical orbit around the Sun takes it from 76 AU out to 937 AU. With such an extended orbit, Sedna requires more than 11,000 years to make a single trip around the Sun. When discovered in 2003, Sedna was about 90 AU from the Sun and getting closer. It will reach its perihelion in 2076. Herschel Space Observatory data suggest an albedo of 0.30 and a size of 1,000 km. Sedna has no known moon, so it is difficult to estimate its mass. Water and methane ices have been detected in its spectrum. Like dwarf planet Eris, Sedna has a highly eccentric orbit. A second object in the inner Oort Cloud, 2012 VP113, was recently detected. Its distance ranges from 80 AU to 452 AU from the Sun, and it is thought to be about half the size of Sedna.

Inner Solar System objects are close enough to the Sun that disturbances external to the Solar System never exert more than a tiny fraction of the gravitational force of the Sun on them. In the distant Oort Cloud, however, comet nuclei are so far from the Sun, and the Sun's gravitational force on them is so weak, that they are barely bound to the Sun at all. The tug of a slowly passing star or interstellar cloud can compete with the Sun's gravity, significantly stirring up the Oort Cloud and changing the orbits of its objects. If the interaction adds to the orbital energy of a comet nucleus, the comet may move outward to an even more distant orbit or perhaps escape from the Sun completely. A comet nucleus that loses orbital energy as a result of this type of interaction will fall inward. Some of these comet nuclei come all the way into the inner Solar System, where they may appear briefly in Earth's skies before returning once again to the Oort Cloud.

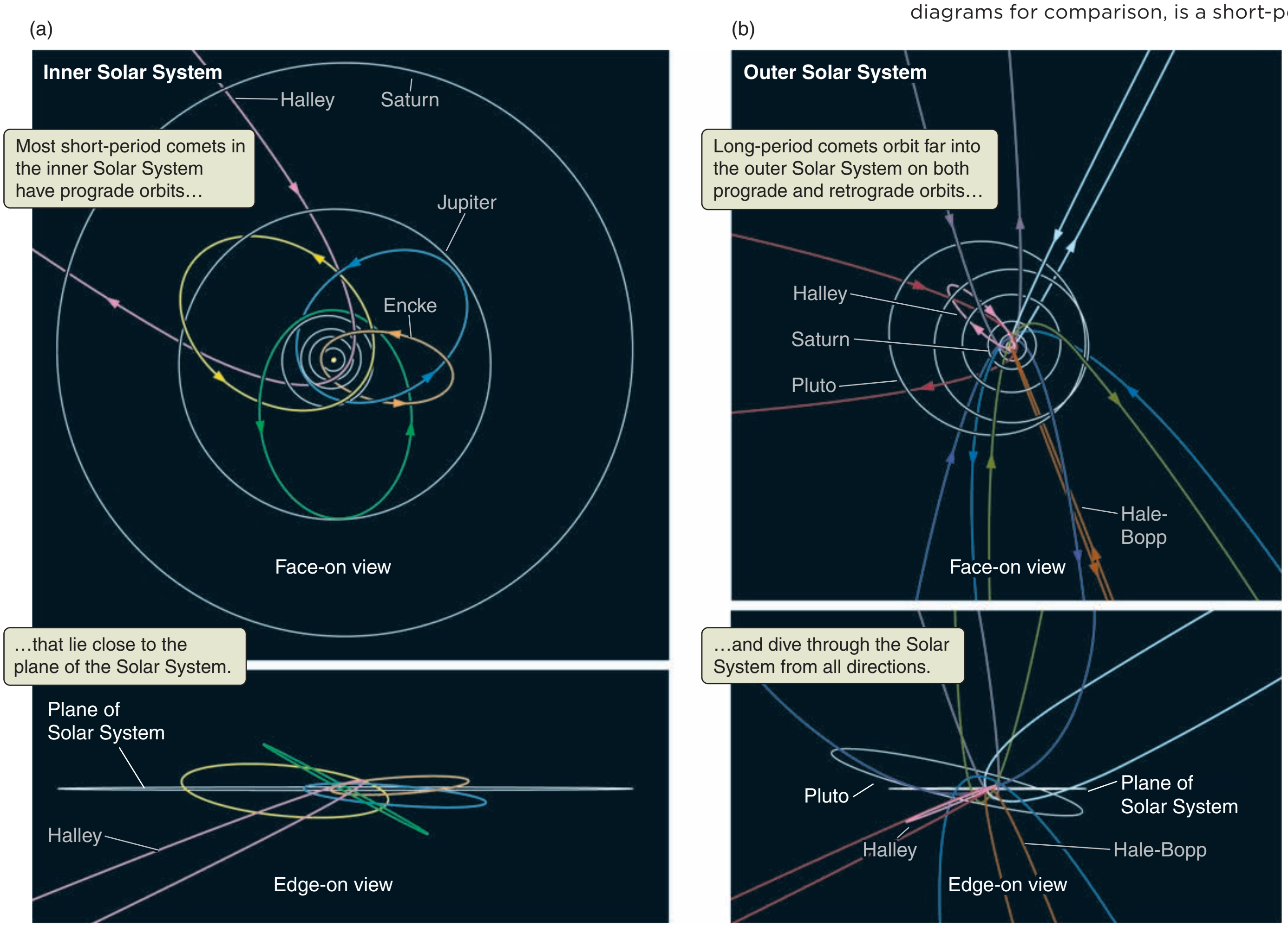
The Orbits of Comets

The lifetime of a comet nucleus depends on how frequently it passes by the Sun and how close it gets. There are about 400 known **short-period comets**, which by definition have periods less than 200 years. Additionally, each year astronomers discover about six new **long-period comets**, whose orbital periods are longer than 200 years. The total number of long-period comets observed to date is about 3,000.

Figure 12.13 shows the orbits of a number of comets, nearly all highly elliptical, with one end of the orbit close to the Sun and the other in the distant parts of the Solar System. Most comets passing through the inner Solar System have long orbital periods that carry them back to the Oort Cloud or the Kuiper Belt. Long-period comets were scattered to the outer Solar System by gravitational interactions, so they come into the inner Solar System from all directions. Some orbit the Sun in the same direction that the planets orbit (prograde), and some orbit in the opposite direction (retrograde).

▶II **AstroTour:** Cometary Orbits

Figure 12.13 This figure illustrates the orbits of a number of comets in face-on and edge-on views of the Solar System. Populations of (a) short-period comets and (b) long-period comets have very different orbital properties. Comet Halley, which appears in both diagrams for comparison, is a short-period comet.



Conversely, short-period comets tend to be prograde and to have orbits in the ecliptic plane, and they frequently pass close enough to a planet for its gravity to change the comet's orbit about the Sun. Short-period comets presumably originated in the Kuiper Belt, but as they fell in toward the Sun, they were forced into their current short-period orbits relatively close to the Sun by gravitational encounters with Jupiter.

Comet Halley is the brightest and most famous of the short-period comets. In 1705, Edmund Halley, using the gravitational laws of his colleague Isaac Newton, noted that a bright comet from 1682 had an orbit remarkably similar to those of comets seen in 1531 and 1607. He concluded that all three were the same comet and predicted that it would return in 1758. When it reappeared, astronomers quickly named it Halley's Comet and heralded it as a triumph for the genius of both Newton and Halley. Comet Halley's highly elongated orbit takes it from perihelion, about halfway between the orbits of Mercury and Venus, out to aphelion beyond the orbit of Neptune. Astronomers and historians have now identified possible sightings of the comet that go back at least to 240 BCE. Comet Halley has an average period of 76 years. Its most recent appearance was in 1986, and it was not especially spectacular compared to 1910 because in 1986, Comet Halley and Earth were on opposite sides of the Sun. Comet Halley will return and become visible to the naked eye once again in summer 2061.

Hundreds of long-period comets have well-determined orbits. Some have orbital periods of hundreds of thousands or even millions of years. Almost all their time is spent in the Oort Cloud in the frigid, outermost regions of the Solar System. Orbits of a few long-period comets are shown in Figure 12.13b. These are the comets that reveal the existence of the Oort Cloud. Because of their very long orbital periods, these comets have made at most one appearance throughout the course of recorded history.

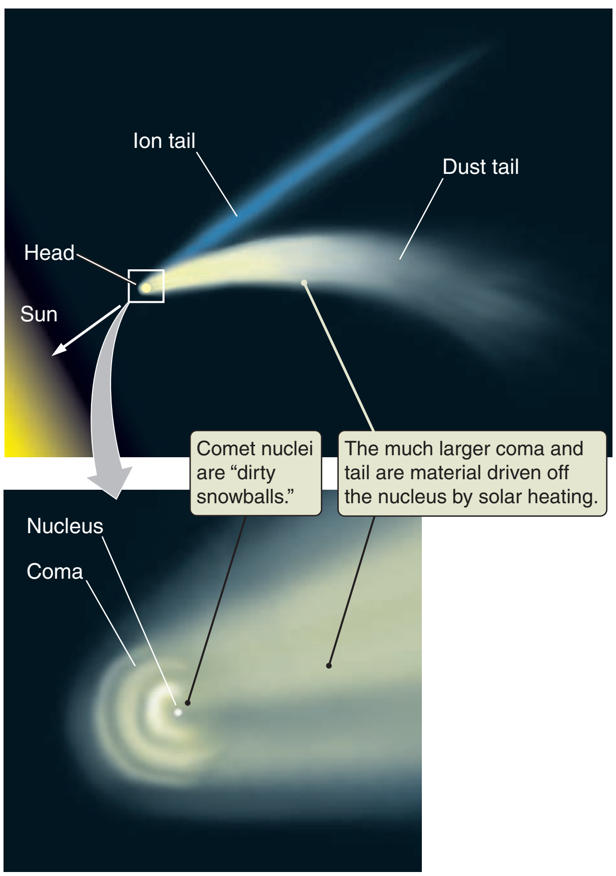


Figure 12.14 The principal components of a fully developed active comet are the nucleus, the coma, and two types of tails called the dust tail and the ion tail. Together, the nucleus and the coma are called the head.

Anatomy of an Active Comet

Unlike asteroids, which have been through a host of chemical and physical changes as a result of collisions, heating, and differentiation, most comet nuclei have been preserved over the past 4.6 billion years by the "deep freeze" of the outer Solar System. Comet nuclei are made of the most nearly pristine material remaining from the formation of the Solar System.

The comet nucleus at the center is the smallest component of a comet, but it is the source of all the material that we see stretched across the sky as the comet nears the Sun (**Figure 12.14**). Comet nuclei range in size from a few dozen meters to several hundred kilometers across. These "dirty snowballs" are composed of ice, organic compounds, and dust grains. They have been described as being similar to deep-fried ice cream, with a soft and porous interior surrounded by a crunchy crust of hardened water-ice crystals, topped off with sooty dust and organic molecules.

As a comet nucleus nears the Sun, sunlight heats its surface, vaporizing ices that stream away from its nucleus, and these gases carry dust particles along with them. This process of conversion from solid to gas is called **sublimation**. For example, dry ice (frozen carbon dioxide) does not melt like water ice but instead turns directly into carbon dioxide gas. Dry ice sublimates—that is why it is called "dry." Set a piece of dry ice out in the Sun on a summer day, and you will get a pretty good idea of what happens to a comet. The gases and dust driven from the nucleus of an active comet form a nearly spherical atmospheric cloud around the nucleus called the

coma. The nucleus and the inner part of the coma are sometimes called the comet's **head**. Pointing from the head of the comet in a direction more or less away from the Sun are long streamers of dust, gas, and ions called **tails**.

The tails are the largest and most spectacular part of a comet. The tails are also the "hair" for which comets are named. (*Comet* comes from the Greek word *kometes*, which means "hairy one.") Active comets have two different types of tails, as shown in Figure 12.14. One is the **ion tail**. Many of the atoms and molecules that make up a comet's coma are ions. Because they are electrically charged, ions in the coma feel the effect of the solar wind—the stream of charged particles that blows continually away from the Sun. The solar wind pushes on these ions, rapidly accelerating them to speeds of more than 100 kilometers per second (km/s)—far greater than the orbital velocity of the comet itself—and sweeps them out into a long wispy structure. Because the particles that make up the ion tail are so quickly picked up by the solar wind, an ion tail is usually very straight: beginning at the head of the comet, an ion tail points directly away from the Sun.

Dust particles in the coma can also have a net electric charge and feel the force of the solar wind. Sunlight also exerts a force on cometary dust. But dust particles are much more massive than individual ions, so they are accelerated more gently and do not reach such high relative speeds as those of the ions. As a result, the dust particles are unable to keep up with the comet, and the **dust tail** often curves away from the head of the comet as the dust particles are gradually pushed from the comet's orbit in the direction away from the Sun (Figure 12.14).

Figure 12.15 shows the tails of a comet at various points in its orbit. Remember that both types of tails always point away from the Sun, regardless of which direction the comet is moving. As the comet approaches the Sun, its two tails trail behind its nucleus. But the tails extend ahead of the nucleus as the comet moves away from the Sun. Tails vary greatly from one comet to another. Some comets display both types of tails simultaneously; others, for reasons that are not understood, produce no tails at all. A tail often forms as a comet crosses the orbit of Mars, where the increase in solar heating drives gas and dust away from the nucleus.

The gas in a comet's tail is even more tenuous than the gas in its coma, with densities of no more than a few hundred particles per cubic centimeter. This is much, much less than the density of Earth's atmosphere, which at sea level contains more than 10^{19} molecules per cubic centimeter. Dust particles in the tail are typically about 1 micron (μm) in diameter, roughly the size of smoke particles.

The nuclei of short-period comets have been badly worn out by their repeated exposure to heating by the Sun. As the volatile ices are driven from a nucleus, some of the dust and organics are left behind on the surface. The buildup of this covering slows down cometary activity. (Envision how, as a pile of dirty snow melts, the dirt left behind is concentrated on the surface of the snow.) In contrast, long-period comets are usually relatively pristine. More of their supply of volatile ices still remains close to the surface of the nucleus, and they can produce a truly magnificent show.

Most naked-eye comets develop first a coma and then an extended tail as they approach the inner Solar System. Comet McNaught in 2007 was such a comet, and it was the brightest to appear in nearly 50 years. The comet's nucleus and coma were visible in broad daylight as its orbit carried it within 25 million km of the Sun. When Comet McNaught had passed behind the Sun and next appeared in the evening skies to observers in the Southern Hemisphere, its tail had grown

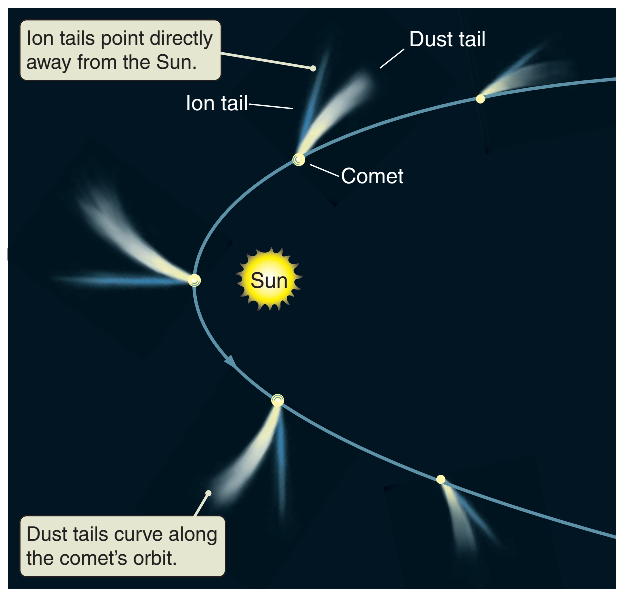


Figure 12.15 This drawing illustrates the orientation of the dust and ion tails at several points in a comet's orbit. The ion tail points directly away from the Sun, while the dust tail curves along the comet's orbit.

G X U **V** I R

Figure 12.16 Comet McNaught in 2007 was the brightest comet to appear in decades, but its true splendor was visible only to observers in the Southern Hemisphere.

to a length of more than 160 million km and stretched 35° across the sky (**Figure 12.16**). Comet McNaught came into the inner Solar System from the Oort Cloud, but it left on a path that will carry it out of the Solar System.

Comet Hale-Bopp in 1997 was a spectacular long-period comet with a long, beautiful tail (**Figure 12.17**). Hale-Bopp was a large comet, with a nucleus estimated at 60 km in diameter. It was discovered far from the Sun, near Jupiter's orbit, 2 years before its perihelion passage. This early discovery extended the total time available to study its development and plan observations as it approached the Sun. Warmed by the Sun, the nucleus produced large quantities of gas and dust and as much as 300 tons of water per second, with lesser amounts of carbon monoxide, sulfur dioxide, cyanogen, and other gases. Comet Hale-Bopp will continue its outward journey for more than 1,000 years, and it will not return to the inner Solar System until sometime around the year 4530.

Comet Ikeya-Seki is a member of a family of comets called **sungrazers**, comets whose perihelia are located very close to the surface of the Sun. Many sungrazers fail to survive even a single orbit of the Sun. Ikeya-Seki became so bright as it neared perihelion in 1965 that it was visible in broad daylight, close to the Sun in the sky. Sungrazers generally come in groups, with successive comets following in nearly identical orbits. Each member of such a group started as part of a single larger nucleus that broke into pieces during an earlier perihelion passage.

A half dozen or so long-period comets arrive each year. Most pass through the inner Solar System at relatively large distances from Earth or the Sun and never become bright enough to attract much public attention. On average, a spectacular comet appears about once per decade.

Visits to Comets

Comets provide an engineering challenge for spacecraft designers. There is seldom enough advance knowledge of a comet's visit or its orbit to mount a successful mission to intercept it. The relative speed between an Earth-launched spacecraft and a comet can be extremely high. Observations must be made very quickly, and there is a danger of high-speed collisions with debris from the nucleus. About a dozen spacecraft have been sent to rendezvous with comets, including five spacecraft sent to Comet Halley by the Soviet, European, and Japanese space agencies in 1986. Much of what we know about comet nuclei and the innermost parts of the coma comes from data sent back by these missions. The spacecraft observed gas and dust jets, impacts craters, and ice and dust on the comet nuclei.

G X U **V** I R

Figure 12.17 Comet Hale-Bopp was a great comet in 1997. The ion tail is blue in this image, and the dust tail is white.

Two Soviet *Vega* and the European *Giotto* spacecraft entered the coma of Comet Halley when they were still nearly 300,000 km from its nucleus. We learned that the dust from Comet Halley was a mixture of light organic substances and heavier rocky material, and the gas was about 80 percent water and 10 percent carbon monoxide with smaller amounts of other organic molecules. The surface of Comet Halley's nucleus is among the darkest known objects in the Solar System, which means that it is rich in complex organic matter that must have been present as dust in the disk around the young Sun—perhaps even in the interstellar cloud from which the Solar System formed. As the three spacecraft passed close by Halley's nucleus, they observed jets of gas and dust moving away from its surface at speeds of up to 1 km/s, far above the escape velocity. By observing the jets of material streaming away from the nucleus of Halley, planetary scientists estimated that Comet Halley must have lost one-tenth of 1 percent of its mass as it went around the Sun.

Several space missions have visited short-period comets. In 2004, NASA's *Stardust* spacecraft flew within 235 km of the nucleus of Comet Wild 2. Comet Wild 2 had previously resided in the region between the orbits of Jupiter and Uranus, but a close encounter with Jupiter in 1974 had perturbed its orbit, bringing this relatively pristine body closer to the Sun as it traveled between the orbits of Jupiter and Earth. At the time of *Stardust*'s encounter with Wild 2, the comet had made only five trips around the Sun in its new orbit. Wild 2's nearly spherical nucleus is about 5 km across. At least 10 gas jets were active, some of which carried large chunks of surface material. The surface of Wild 2 is covered with features that may be impact craters modified by ice sublimation, small landslides, and erosion by jetting gas (**Figure 12.18**). Some craters show flat floors, suggesting a relatively solid interior beneath a porous surface layer.

The *Stardust* mission collected dust samples from Wild 2, which were returned to Earth in 2006. It found new kinds of organic materials unlike any seen before in materials from space. They are more primitive than those observed in meteorites and may have formed before the Solar System itself. These grains can be used to investigate the conditions under which the Sun and planets formed. Minerals that form at high temperature have also been found, supporting the idea that the solar wind blew material out of the inner Solar System very early in the system's history. Scientists will be studying the particles from this mission in detail for many years.

In 2005, NASA's *Deep Impact* spacecraft launched a 370-kg impacting projectile into the nucleus of Comet Tempel 1 at a speed of more than 10 km/s. The impact sent 10,000 tons of water and dust flying off into space at speeds of 50 meters per second (m/s—**Figure 12.19**). A camera mounted on the projectile snapped photos of its target until it was vaporized by the impact. Observations of the event were made both locally by *Deep Impact* and back on Earth by orbiting and ground-based telescopes. Water, carbon dioxide, hydrogen cyanide, iron-bearing minerals, and a host of complex organic molecules were identified in the Comet Tempel 1 impact. The comet's outer layer is composed of fine dust with a consistency of talcum powder. Beneath the dust are layers made up of water ice and organic materials. Well-formed impact craters, which had been absent in close-up images of Comet Wild 2, were also seen.

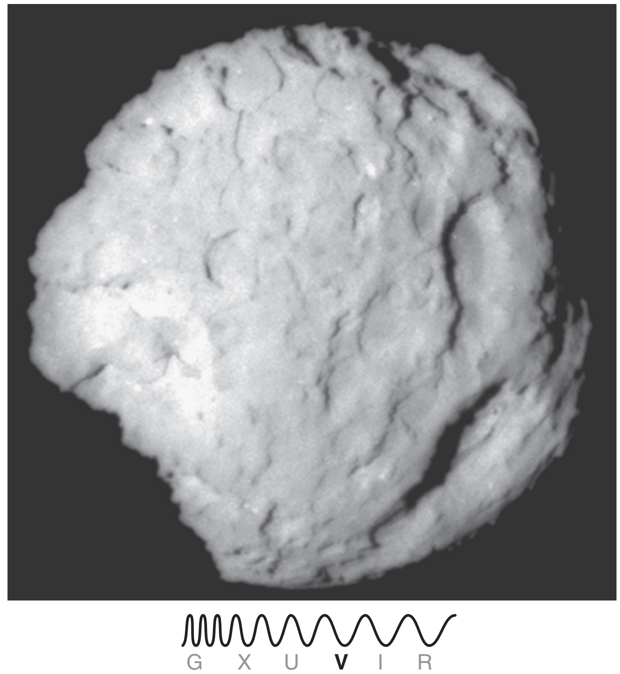


Figure 12.18 The nucleus of Comet Wild 2 was imaged by the *Stardust* spacecraft, which also sampled its tail.

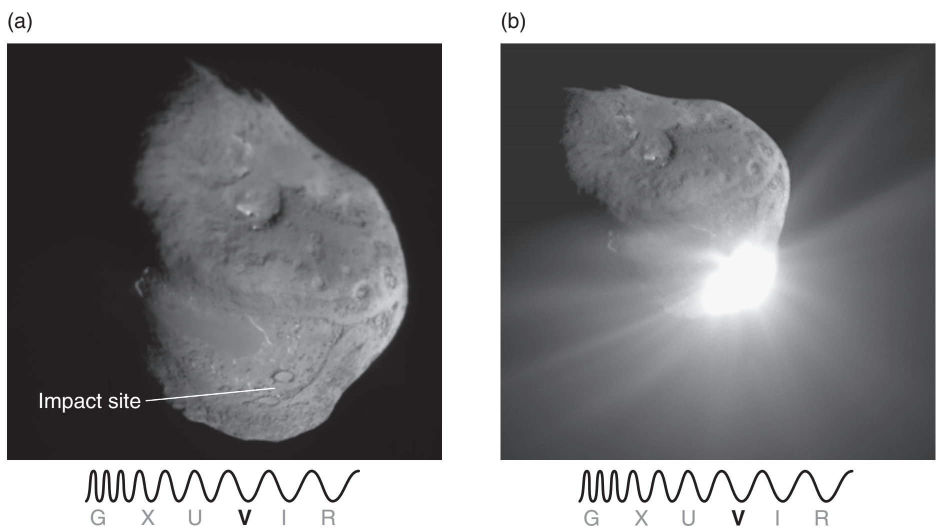


Figure 12.19 (a) The surface of the nucleus of Comet Tempel 1 is shown just before impact by the *Deep Impact* projectile. The impact occurred between the two 370-meter-diameter craters located near the bottom of the image. The smallest features appearing in this image are about 5 meters across. (b) Sixteen seconds after the impactor struck the comet, the parent spacecraft took this image of the initial ejecta.

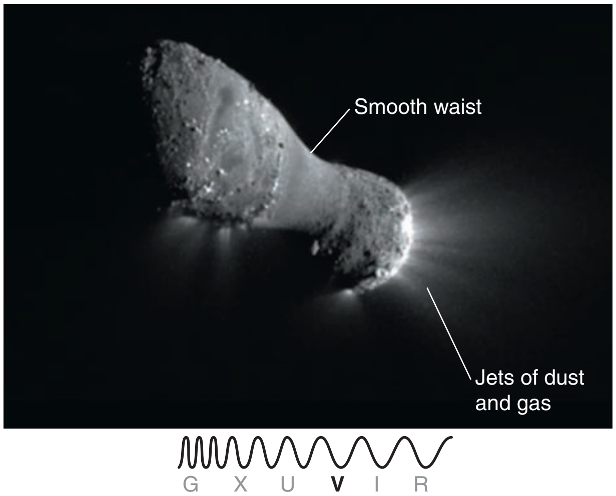


Figure 12.20 This image of Comet Hartley 2 taken by the *EPOXI* spacecraft reveals two distinct surface types. Water seeps through the dust at the smooth "waist" of the comet's nucleus, whereas carbon dioxide jets shoot gas, dust, and chunks of ice from the rough areas.

In 2010, the *EPOXI* spacecraft flew past Comet Hartley 2 (**Figure 12.20**), imaging not only jets of dust and gas, indicating a remarkably active surface, but also an unusual separation of rough and smooth areas. The narrow part at the middle of the image (Figure 12.20) is a smooth inactive area where ejected material has fallen back onto the cometary nucleus. Carbon dioxide jets shoot out from the rough areas. Further observations by the Herschel Space Observatory showed that the water on this comet has the same ratio of hydrogen isotopes as that of the water in Earth's oceans. This suggests that some of Earth's water could have originated in the Kuiper Belt. Measurements of water on comets from the Oort Cloud have a different ratio, and so they have been ruled out as the source of Earth's water.

The European Space Agency spacecraft *Rosetta* visited Comet 67P/Churyumov-Gerasimenko in 2014 (see the chapter-opening figure). A small separate spacecraft soft-landed on the comet and sent back data for 2.5 days before running out of power. The landing site was dust-covered solid ice, too thick to drill. *Rosetta* orbited the comet as it approached the Sun and observed changes as it heated up. Dust and gas were released from the comet, including a bright jet.

CHECK YOUR UNDERSTANDING 12.3

The nucleus of a comet is mostly: (a) solid ice; (b) solid rock; (c) a porous mix of ice and dust; (d) frozen carbon dioxide.

12.4 Meteorites Are Remnants of the Early Solar System

Nebraska Simulation: Driving through Snow

Comet nuclei that enter the inner Solar System generally disintegrate within a few hundred thousand years as a result of their repeated passages near the Sun. Asteroids have much longer lives but still are slowly broken into pieces from occasional collisions with each other. The disintegration of comet nuclei and collisions between asteroids create most of the debris that fills the inner part of the Solar System. As Earth and other planets move along in their orbits, they continually sweep up this fine debris. The cometary and asteroidal debris is the source of most of the *meteoroids* that Earth encounters. Meteoroids are small solid bodies ranging in size from 10 μm to 100 meters. When a meteoroid enters Earth's atmosphere, frictional heat causes the air to glow, producing an atmospheric phenomenon called a *meteor*. If a meteoroid survives to reach the planet's surface, we call it a *meteorite*. Earth sweeps up some 100,000 kg of meteoritic debris every day, and particles smaller than 100 μm eventually settle to the ground as fine dust. In this section, we will look more closely at meteorites and what can be learned about the early Solar System from them.

Observations of Meteors

If you stand outside for a few minutes on a moonless, starry night, away from bright city lights, you will almost certainly see a meteor, commonly known as a *shooting star*. The larger pieces that survive the plunge through Earth's atmosphere are usually fragments of asteroids. Most of the smaller pieces that burn up in the atmosphere before reaching the ground are cometary fragments typically less than a centimeter across and having about the same density as cigarette ash.

A 1-gram meteoroid (about half the mass of a dime) entering Earth's atmosphere at 50 km/s has a kinetic energy comparable to that of an automobile cruising along at the fastest highway speeds. Scientists measuring meteor heights with radar find that the altitudes of meteors are between 50 and 150 km. Most meteoroids are so small and fragile that they burn up completely before reaching Earth's surface. A meteor may streak across 100 km of Earth's atmosphere and last at most a few seconds. Meteorites likely litter the surfaces of all solid planets and moons.

Nearly all ancient cultures were fascinated by these rocks from the sky. Iron from meteorites was used to make the earliest tools. Early Egyptians preserved meteorites along with the remains of their pharaohs, Japanese placed them in Shinto shrines, and ancient Greeks worshipped them. Despite numerous eyewitness accounts of meteorite falls, however, many people were slow to accept that these peculiar rocks actually come from far beyond Earth. By the early 1800s, scientists had documented so many meteorite falls that their true origin was indisputable. Today, hardly a year passes without a recorded meteorite fall, including some that have caused damage.

Fragments of asteroids are much denser than cometary meteoroids. If an asteroid fragment is large enough—about the size of your fist—it can survive all the way to the ground to become a meteorite. The fall of a 10-kg meteoroid can produce a fireball so bright that it lights up the night sky more brilliantly than the full Moon. Such a large meteoroid, traveling many times faster than the speed of sound, may create a sonic boom heard hundreds of kilometers away. It may even explode into multiple fragments as it nears the end of its flight. Some fireballs glow with a brilliant green color, caused by elements in the meteoroid that created them.

Meteor showers occur when Earth's orbit crosses the orbit of a comet or asteroid and passes through a concentration of cometary or asteroidal debris. During a shower, many meteors can be observed in just a few hours. More than a dozen comets and at least two asteroids have orbits that come close enough to Earth's orbit to produce annual meteor showers, as listed in **Table 12.1**. Because the meteoroids in a shower are all in similar orbits, they all enter Earth's atmosphere moving in the same direction—the paths through the sky are parallel to one another. Therefore, all the meteors appear to originate from the same point in the sky (**Figure 12.21a**),

TABLE 12.1 Selected Meteor Showers

Shower	Approximate Date	Parent Object
Quadrantids	January 3–4	Asteroid 2003 EH1
Lyrids	April 21–22	Comet Thatcher
Eta Aquariids	May 5–6	Comet Halley
Perseids	August 12–13	Comet Swift-Tuttle
Draconids	October 8–9	Comet Giacobini-Zinner
Orionids	October 21–22	Comet Halley
Taurids	November 5–6	Comet Encke
Leonids	November 17–18	Comet Tempel-Tuttle
Geminids	December 13–14	Asteroid Phaethon
Ursids	December 22–23	Comet Tuttle

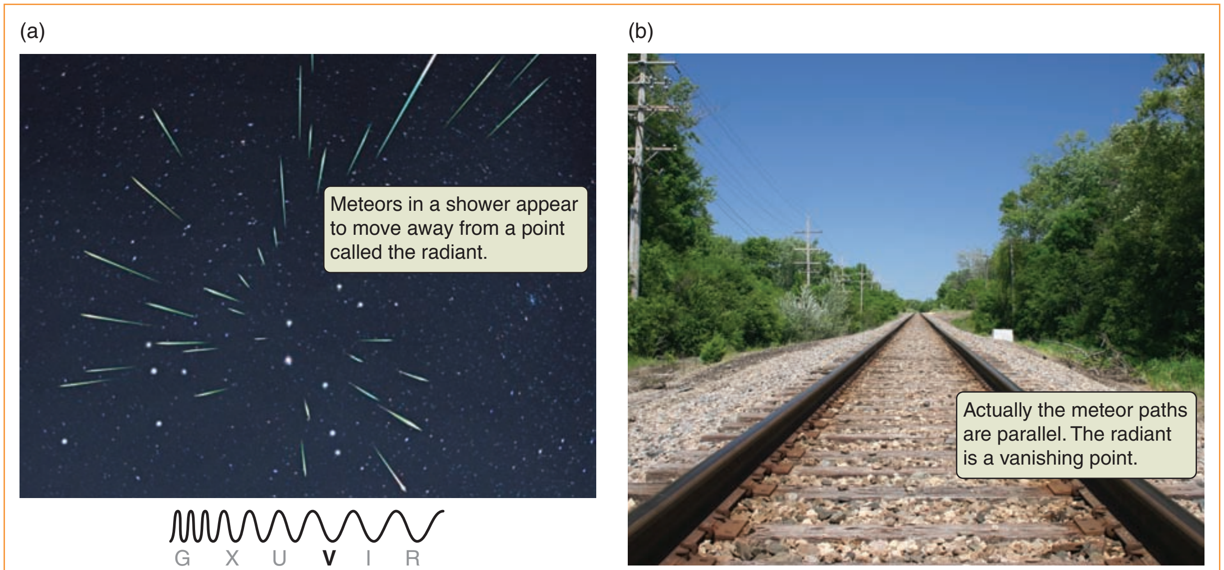


Figure 12.21 (a) Meteors appear to stream away from the radiant of the Leonid meteor shower. (b) Such streaks are actually parallel paths that appear to emerge from a vanishing point, as in our view of these railroad tracks.

VISUAL ANALOGY

just as the parallel rails of a railroad track appear to vanish to a single point in the distance (Figure 12.21b). This point is called the shower's **radiant**.

For example, the Perseid shower in August is the result of Earth crossing the orbit of Comet Swift-Tuttle. Although spread out along the comet's orbit, the debris is more concentrated in the vicinity of the comet itself. In 1992, Comet Swift-Tuttle returned to the inner Solar System for the first time since its discovery in 1862, resulting in an exceptional Perseid meteor shower with counts of up to 500 meteors per hour.

In mid-November of each year, Earth passes almost directly through the orbit of Comet Tempel-Tuttle, a short-period comet with an orbital period of 33.2 years. This produces the Leonid meteor shower, which is usually weak because in most years, Comet Tempel-Tuttle distributes little of its debris around its orbit. In 1833 and 1866, however, Tempel-Tuttle was not far away when Earth passed through its orbit, and the Leonid showers were so intense that meteors filled the sky with as many as 100,000 meteors per hour. Further perturbations of the comet's orbit caused a spectacular Leonid shower in 1966 that may have produced as many as a half-million meteors per hour. The Leonid shower put on less spectacular but still impressive shows between 1999 and 2003, when several thousand meteors per hour were seen.

Types of Meteoroids

As asteroids orbit the Sun, they occasionally collide with each other, chipping off smaller rocks and bits of dust. Sometimes, one of these fragments is captured by Earth's gravity and survives its fiery descent through Earth's atmosphere as a meteor. Thousands of meteorites reach the surface of Earth every day, but only a tiny fraction of these are ever found and identified. Antarctica offers the best meteorite hunting in the world because in many places, the only stones to be found on the ice are meteorites. Because Antarctica is actually very dry, Antarctic meteorites also tend to show little weathering or contamination from terrestrial dust or organic compounds, which makes them excellent specimens for study.

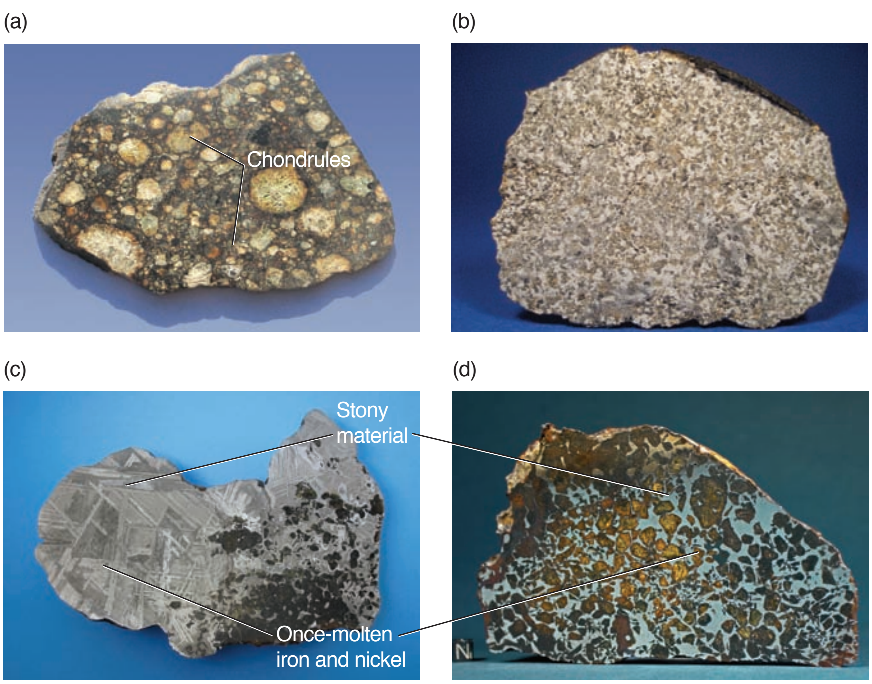


Figure 12.22 Cross sections of several kinds of meteorites: (a) a chondrite (a stony meteorite with chondrules); (b) an achondrite (a stony meteorite without chondrules); (c) an iron meteorite; (d) a stony-iron meteorite.

Meteorites are grouped into three categories according to their materials and the degree of differentiation they experienced within their parent bodies (**Figure 12.22**). More than 90 percent of meteorites are included in the first category, **stony meteorites**, which are similar to terrestrial silicate rocks. A stony meteorite is characterized by the thin coating of melted rock that forms as it passes through the atmosphere. Many stony meteorites contain small round spherules called **chondrules**, once-molten droplets that rapidly cooled to form crystallized spheres ranging in size from that of sand grains to that of marbles. Stony meteorites containing chondrules are called **chondrites** (Figure 12.22a); those without chondrules are known as **achondrites** (Figure 12.22b). **Carbonaceous chondrites** are chondrites that are rich in carbon: these are the most primitive of the meteorites. Indirect measurements suggest that these meteorites are about 4.56 billion years old—consistent with all other measurements of the time that has passed since the Solar System was formed.

The second major category of meteorites, **iron meteorites** (Figure 12.22c), comes from M-type asteroids. Iron meteorites can be recognized by their melted and pitted appearance generated by frictional heating as it streaked through the atmosphere. Many iron meteorites are never found, either because they land in water or because they are not recognized as meteorites. The Mars *Opportunity* rover discovered a few iron meteorites on the martian surface (**Figure 12.23**). Both their appearance—typical of iron meteorites found on Earth—and their position on the smooth, featureless plains made them instantly recognizable.

The third category of meteorites is the **stony-iron meteorites**, which consist of a mixture of rocky material and iron-nickel alloys (Figure 12.22d). Stony-iron meteorites are relatively rare.

Figure 12.23 This 7-foot-long iron meteorite lying on the surface of Mars was imaged by the Mars exploration rover *Curiosity*.

Meteorites and the History of the Solar System

Meteorites are extremely valuable because they are samples of the same relatively pristine material that makes up asteroids. Astronomers can take meteorites into the laboratory and study them. Scientists compare meteorites to rocks found on Earth and the Moon and contrast their structure and chemical makeup with rocks studied by spacecraft that have landed on Mars and Venus. Comparing the spectra of meteorites with those of asteroids and planets reveals their origin.

Meteorites come from asteroids, which derive from stony-iron planetesimals. A few planetesimals in the region between the orbits of Mars and Jupiter evolved toward becoming tiny planets before being shattered by collisions. Some became volcanically active, with eruption of lava onto their surfaces. But rather than forming planets, these planetesimals broke into pieces in collisions with other planetesimals.

Some types of meteorites fail to follow the patterns just discussed. Whereas most achondrites have ages in the range 4.5 billion to 4.6 billion years, some are less than 1.3 billion years old. Other achondrites are chemically and physically similar to the soil and the atmospheric gases that NASA's lander instruments have measured on Mars. The similarities are so strong that most planetary scientists think these meteorites are pieces of Mars that were knocked into space by large asteroidal impacts—so that researchers can study pieces of Mars in laboratories here on Earth. In 1996, a NASA research team announced that the meteorite ALH84001, found in Antarctica, showed possible physical and chemical evidence of past life on Mars, but the claim is still debated (see Chapter 24).

Another group of meteorites bear striking similarities to samples returned from the Moon. Like the meteorites from Mars, these are chunks of the Moon that were blasted into space by impacts and later fell to Earth. It is possible, therefore, that meteorites from Earth fell on the Moon. If they are ever collected from the unchanging Moon, they could tell us about what conditions were like on the early Earth.

Zodiacal Dust

Like meteoroids, **zodiacal dust** is a mixture of cometary debris and ground-up asteroidal material. Just as you can "see" sunlight streaming through an open window by observing its reflection from dust drifting in the air, you can see the sunlight reflected off tiny zodiacal dust particles that fill the inner parts of the Solar System close to the plane of the ecliptic. On a clear, moonless night, not long after the western sky has grown dark, this dust is visible as a faint column of

Figure 12.24 Zodiacal light shines in the western sky after sunset as seen from the La Silla Observatory in Chile.

light slanting upward from the western horizon along the path of the ecliptic. This band, called the **zodiacal light**, can also be seen in the eastern sky just before dawn (**Figure 12.24**). With good eyes and an especially dark night, you may be able to follow the zodiacal dust band all the way across the sky. In its brightest parts, the zodiacal light can be several times brighter than the Milky Way, for which it is sometimes mistaken.

The dust grains are roughly a millionth of a meter in diameter—the size of smoke particles. In the vicinity of Earth, each cubic kilometer of space contains only a few particles of zodiacal dust. The total amount of zodiacal dust in the entire Solar System is estimated to be 10^{16} kg, equivalent to a solid body about 25 km across, or roughly the size of a large comet nucleus. Grains of zodiacal dust are constantly being lost as they are swept up by planets or pushed out of the Solar System by the pressure of sunlight. Such interplanetary dust grains have been recovered from Earth's upper atmosphere by aircraft flying very high. If not replaced by new dust from comets, all zodiacal dust would be gone within the brief span of 50,000 years.

In the infrared region of the spectrum, thermal emission from the band of warm zodiacal dust makes it one of the brightest features in the sky. It is so bright that astronomers wanting to observe faint infrared sources are frequently hindered by its foreground glow.

CHECK YOUR UNDERSTANDING 12.4

Meteorites contain clues to which of the following? (Choose all that apply.) (a) the age of the Solar System; (b) the temperature in the early solar nebula; (c) changes in the composition of the primitive Solar System; (d) changes in the rate of cratering in the early Solar System; (e) the physical processes that controlled the formation of the Solar System.

12.5 Collisions Still Happen Today

Almost all hard-surfaced objects in the Solar System still bear the scars of a time when tremendous impact events were common. Although such impacts are far less frequent today than they once were, they still happen. In this section, we'll examine a few examples of recent collisions.

Comet Shoemaker-Levy 9 Collided with Jupiter

Early in the 20th century, the orbit of a comet nucleus called Shoemaker-Levy 9 from the Kuiper Belt was perturbed, and the comet's new orbit carried it close to Jupiter. Eventually, it was captured by Jupiter and orbited the planet. In 1992, this comet passed so close to Jupiter that tidal stresses broke it into two dozen major fragments, which subsequently spread out along its orbit. The fragments took one more 2-year orbit around the planet, and throughout a week in 1994, the entire string of fragments crashed into Jupiter. The impacts occurred just behind the limb of the planet, so they were not visible on Earth until Jupiter's rotation put the impact points in view. Astronomers using ground-based telescopes and the Hubble Space Telescope could see immense plumes rising from the impacts to heights of more than 3,000 km above the cloud tops at the limb. The debris in these plumes then rained back onto Jupiter's stratosphere, causing ripples like pebbles thrown into a pond. **Figure 12.25** shows some HST images of the impact features. Sulfur and carbon compounds released by the impacts formed Earth-sized scars in the atmosphere that persisted for months.

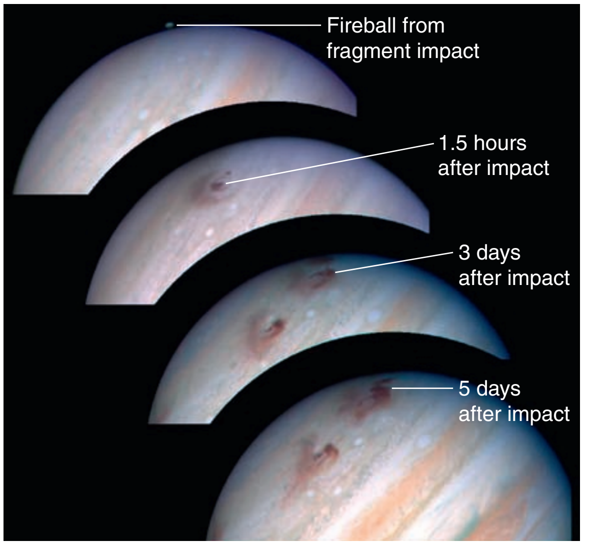


Figure 12.25 HST images of the evolution of the scar left by one fragment of Comet Shoemaker-Levy 9 when it impacted Jupiter in 2004.

Collisions with Earth

In summer 1908, a remote region of western Siberia was blasted with the energy equivalent of 2,000 times the energy of the atomic bomb dropped on Hiroshima. Eyewitness accounts detailed the destruction of dwellings, the incineration of reindeer (including one herd of 700), and the deaths of at least five people. Although trees were burned or flattened over more than 2,150 square kilometers (km^2)—an area greater than metropolitan New York City—no crater was left behind. The Tunguska event (named for the nearby river) was the result of a tremendous high-altitude explosion that occurred when a small body hit Earth's atmosphere, ripped apart, and formed a fireball before reaching Earth's surface. Recent expeditions to the Tunguska area have recovered resin from the trees blasted by the event. Chemical traces in the resin suggest that the impacting object may have been a stony asteroid.

In February 2013, a known near-Earth object about half the size of an American football field passed so close to Earth that it came within the orbit of man-made satellites. This near miss was uneventful, and the object simply continued on its way. However, in an unrelated event on the same day, a previously unknown meteoroid estimated to have a radius of about 20 meters exploded over Chelyabinsk, Russia. The shock waves from this explosion damaged thousands of buildings in six cities and injured more than 1,000 people. This was likely the largest impact on Earth since the Tunguska event, and there were many recorded observations of the effect on this less remote location.

From car dashboard, cell phone, and security camera video and images (**Figure 12.26a**), as well as reports of the time between the brighter-than-the-Sun flash and the sonic boom that followed, scientists determined the trajectory and speed of the incoming object as it traveled through the atmosphere. They estimate a preimpact orbit of the object in the inner asteroid belt and think it originally broke off from a known 2-km-sized asteroid. From small pieces collected over a wide area and from a large, 600-kg chunk found in a frozen-over lake (Figure 12.26b), scientists could analyze the object's composition and density. It seems to

(a)

(b)

Figure 12.26 (a) In February 2013, a meteoroid entered the atmosphere over Russia, creating a fireball that eyewitnesses said was brighter than the Sun. (b) A 600-kg piece of the meteorite on display.

12.2 Working It Out Impact Energy

How much energy can be released from the impact of a comet nucleus? The kinetic energy of a moving object is given by

$$E_K = \frac{1}{2}mv^2$$

where E_K is the kinetic energy in joules (J), m is the mass in kilograms, and v is the speed in meters per second.

Suppose an asteroid or comet nucleus that is 10 km in diameter with a mass of 5×10^{14} kg hits Earth at a speed of 20 km/s = 20×10^3 m/s. Putting these values into the preceding equation gives us

$$E_K = \frac{1}{2} \times (5 \times 10^{14}\ \text{kg}) \times (20 \times 10^3\ \text{m/s})^2$$

$$E_K = 1.0 \times 10^{23}\ \text{J}$$

How much energy is this? A 1-megaton hydrogen bomb (67 times as energetic as the Hiroshima atomic bomb) releases 4.2×10^{15} J. If we divide the energy from our comet impact by this number, we have

$$\frac{1.0 \times 10^{23}\ \text{J}}{4.2 \times 10^{15}\ \text{J/megaton H-bomb}} = 2.4 \times 10^7$$

$$= 24 \text{ million 1-megaton H-bombs}$$

That's a lot of energy, and it shows why impacts have been so important in the history of the Solar System.

be similar in composition to Itokawa, the asteroid whose dust was collected and returned to Earth. It is estimated that only a few hundredths of 1 percent of the original 10-million-kg mass has been found on the surface of Earth. The energy of the explosion was about 30 times the blast power of a World War II atomic bomb. However, most of that energy went into the atmosphere, heating and breaking up the meteoroid at a much higher altitude than where bombs are detonated, so the effects on the ground were less than that of a bomb.

These impacts are sobering events. The distribution of relatively large asteroids in the inner Solar System indicates it is highly improbable that an asteroid will impact a populated area on Earth within your lifetime. However, there are many comets and smaller asteroids with unknown orbits, and several previously unknown long-period comets enter the inner Solar System each year. If on a collision course with Earth, even a large comet might not be noticed until just a few weeks or months before impact. For example, Comet Hyakutake was discovered only 2 months before it passed near Earth, and a potentially destructive asteroid that just missed Earth in 2002 was not discovered until 3 days *after* its closest approach. The Chelyabinsk meteoroid was in an orbit such that it wasn't detectable at all. Earth's geological and historical record suggests that actual impacts by large bodies are infrequent events.

There may be as many as 10 million asteroids larger than a kilometer across, but only about 130,000 have well-known orbits, and most of the unknown asteroids are too small to see until they come very close to Earth. The U.S. government—along with the governments of several other nations—is aware of the risk posed by Near-Earth Objects. Although the probability of a collision between a small asteroid and Earth is quite small, the consequences could be catastrophic (**Working It Out 12.2**), so NASA has been given a congressional mandate to catalog all NEOs and to scan the skies for those that remain undiscovered.

CHECK YOUR UNDERSTANDING 12.5

How do astronomers determine the origin of a meteorite that reaches Earth?

Origins

Comets, Asteroids, Meteoroids, and Life

Water is essential to life on Earth, so it seems important to know where the water came from. However, the origin of the water is still under debate. Scientists have thought that some of Earth's water was contributed during impacts of icy planetesimals early in the history of the Solar System. The icy planetesimals condensed from the protoplanetary disk surrounding the young Sun and grew to their current size near the orbits of the giant planets. These planetesimals subsequently suffered strong orbital disturbances from the giant planets, which may themselves have been migrating to and from the inner Solar System. In such interactions, some of the planetesimals were flung outward to form the Kuiper Belt and Oort Cloud, and some were thrown inward toward the Sun, possibly hitting Earth. Because much of the mass in comet nuclei and some in asteroids appears to be in the form of water ice, it is possible that some of Earth's current water supply came from this early bombardment. Spacecraft have measured the type of water in several comets and asteroids, and to date the water on Earth best matches that of the primitive carbonaceous chondrites and a few, but not most, comets.

Comets and asteroids can threaten life on Earth. Occasional collisions of comet nuclei and asteroids with Earth have probably resulted in widespread devastation of Earth's ecosystem and in the extinction of many species. Passing stars or the periodic passage of the Sun through giant gas clouds located in denser regions of the Milky Way Galaxy may have resulted in showers of comet nuclei into the inner Solar System, possibly contributing to a change in the climate and mass extinctions. Although such events certainly qualify as global disasters for the plants and animals alive at the time, they also represent global opportunities for new life-forms to evolve and fill the niches left by species that did not survive. As noted in Chapter 8, such a collision with an asteroid or comet likely played a central role in ending the 180-million-year reign of dinosaurs and provided an opportunity for the evolution of mammals.

In studying comets, astronomers may also have found a key to the chemical origins of life on Earth. Comets are rich in complex organic material—the chemical basis for terrestrial life—and cometary impacts on the young Earth may have played a role in chemically seeding the planet. If comets are pristine samples of the material from which the Sun and planets formed, then organic material must be widely distributed throughout interstellar space. Radio telescope observations of vast interstellar clouds throughout the Milky Way confirm the presence of organic material. The fact that asteroid belts and storms of comets have been observed in distant solar systems could have significant implications as astronomers consider the possibility of life elsewhere in the universe.

READING ASTRONOMY NEWS

ARTICLES QUESTIONS

Scientists report on the water observed on Comet 67P/Churyumov-Gerasimenko.

Rosetta Spacecraft Finds Water on Earth Didn't Come from Comets

By **REBECCA JACOBSON, PBS NewsHour**

It's a mystery that has baffled scientists for decades: Where did Earth's water come from?

Some scientists believed comets might have been the original source of the Earth's oceans. But a study published this week in the journal *Science* is sending scientists back to the drawing board. In its first published scientific data, the ROSINA mass spectrometer on board the *Rosetta* probe found that water on Comet 67P/Churyumov-Gerasimenko doesn't match the water on Earth.

The result is surprising, says Kathrin Altwegg, principal investigator for ROSINA at the University of Bern and one of the authors of the study. For decades, scientists had ruled out comets from the Oort Cloud at the very edge of our Solar System as the source of Earth's water.

But three years ago, an analysis of water on the Hartley 2 comet near Jupiter found a perfect match to the Earth's oceans. That finding led scientists to believe that Earth's water could have come from much closer comets, either near Jupiter or in the Kuiper Belt just beyond Neptune. Comet 67P/Churyumov-Gerasimenko is one of those Jupiter family comets, which scientists believe originated in the Kuiper Belt.

"That was a big surprise, but now we are back to what I expected," she said. "I think it's very nice to see the diversity we have in Kuiper Belt, to see that not everything is as simple as it seemed."

To find the origin of Earth's water, scientists analyze the water's "fingerprint," says Claudia Alexander, project scientist of the U.S. *Rosetta* Project at NASA's Jet Propulsion Laboratory. Water has a chemical isotopic signature, which works just like a fingerprint. Planets, comets, even minerals all have a fingerprint, Alexander says, and scientists are looking for a match to Earth's.

On Earth, water is mostly two parts hydrogen and one part oxygen—H_2O. But there's also "heavy" water, Alexander explained, which is made with deuterium—a hydrogen atom with a neutron. That heavy water is what the *Rosetta* spacecraft found on Comet 67P/Churyumov-Gerasimenko. It's also a closer match to the water scientists have found on other comets, ruling them out as Earth's water source, Alexander said.

"The clues don't quite all add up," she said.

Altwegg agrees, saying that it's not likely the other Kuiper Belt comets have a match to Earth's water, but further studies would be helpful.

"You would have to assume that 67P is the exception in the Kuiper Belt," she said. "We need more missions to Kuiper Belt comets, which would be fabulous."

There are several ideas to explain the origin of Earth's water, Alexander said. Some believe that water has been on Earth since its formation, that it was beaten out of other minerals as the planet formed. Others think "wet planetesimals" near Jupiter—which were like planetary Silly Putty, loose sticky blobs of rock and ice, Alexander said—collided with Earth in the early formation of the Solar System.

Alexander believes the answer could be a combination of any of these ideas. The *Dawn* mission in 2015 will study the water on the asteroid Ceres, near Jupiter. If it's a match for Earth's water, it may be another clue, Alexander said. But the *Rosetta* finding is a huge step in solving the mystery, she said.

"I think this is a big deal. . . . For me, I've not always been a believer in the story that comets brought the water," Alexander said. "In some respects, I'm somewhat relieved (this finding) doesn't confirm it. It's more complicated than that. I think we need more forensic evidence to settle the score."

ARTICLES QUESTIONS

1. This article uses the word "fingerprint" when discussing isotope ratios. Earlier in this book, we used it in our discussion of spectral lines. How are the two phenomena similar?
2. What might explain why the water on Comet Hartley 2 matched Earth's water but the water on Comet 67P/Churyumov-Gerasimenko did not?
3. Is the water on other comets more like the water on Comet Hartley 2 or on Comet 67P/Churyumov-Gerasimenko?
4. What are other possible sources of Earth's water?
5. Do an internet search to see if there have been any results from the Dawn mission indicating the type of water observed on Ceres.

Summary

The story of how planetesimals, asteroids, and meteorites are related is a great success of planetary science. Scientists have assembled a wealth of information about this diverse collection of objects to piece together a picture of how planetesimals grow, differentiate, and then shatter in subsequent collisions. This story fits well with the even larger story of how most planetesimals were accreted into the planets and their moons. Comets, asteroids and meteoroids may have supplied Earth with water, volatiles, and organic material in the early history of the Solar System. Impacts with large asteroids or comet nuclei may have led to mass extinctions on Earth that eventually enabled mammals to evolve. Recent spacecraft missions to comets and asteroids have begun to reveal details about the composition of these bodies.

LG 1 **List the categories of small bodies and identify their locations in the Solar System.** Small bodies in the Solar System that orbit the Sun include dwarf planets, asteroids, comets, Kuiper Belt objects, and meteoroids. Most asteroids are located in the main asteroid belt between the orbits of Mars and Jupiter. Comets are small, icy planetesimals that reside in the frigid regions of the Kuiper Belt and the Oort Cloud, beyond the planets. The orbits of nearly all comets are highly elliptical, with one end of the orbit close to the Sun and the other in the distant parts of the Solar System.

LG 2 **Describe the defining characteristics of the dwarf planets in the Solar System.** Pluto, Eris, Haumea, Makemake, and Ceres are classified as dwarf planets because, although they are sufficiently massive to have pulled themselves into round shapes, they are not massive enough to have cleared their surroundings of other bodies and are therefore not planets.

LG 3 **Describe the origin of the different types of asteroids, comets, and meteorites.** Asteroids are small Solar System bodies made of rock and metal. Although early collisions between these planetesimals created several bodies large enough to differentiate, Jupiter's tidal disruption (and possible migration) prevented them from forming a single planet. Comets that venture into the inner Solar System are warmed by the Sun, producing an atmospheric coma and a tail. Meteoroids are small fragments of asteroids and comets. When a meteoroid enters Earth's atmosphere, frictional heat causes the air to glow, producing a meteor. Meteor showers occur when Earth passes through a trail of cometary debris. A meteoroid that survives to a planet's surface is called a meteorite. The various types of meteoroids that are formed depend on the differentiation of the parent body.

LG 4 **Explain how asteroids, comets, and meteoroids provide important clues about the history and formation of the Solar System.** Asteroids, comets, and meteoroids are leftover debris from the formation of the Solar System. Asteroids are composed of the same type of material that became the inner planets, and comets are composed of the same type of material that became the outer planets. They provide samples of the initial composition and properties of the Solar System and furnish samples of material from its entire history.

LG 5 **Describe what has been learned from observations of recent impacts in the Solar System.** Impacts by comets and meteoroids have been observed recently on Jupiter and Earth. The debris from these impacts helps astronomers understand the conditions in the early Solar System when impacts were much more frequent.

? UNANSWERED QUESTIONS

- What can asteroids and comets reveal about the dynamics of the early Solar System? It is not accidental that the main asteroid belt and the Kuiper Belt straddle the orbits of the giant planets. As noted in Chapter 10, the giant planets may have moved around quite a bit in the early Solar System. This migration of giant planets may explain the spread of orbits of the objects in the main asteroid belt. Migration may also have brought icy objects—such as comet nuclei—out of the Kuiper Belt and into the main asteroid belt. The Kuiper Belt would have been closer to the Sun originally, and Jupiter and Saturn would have pushed it outward. And finally, the migration might have sent objects from both belts into the inner Solar System, creating the heavy bombardment of 4 billion years ago.
- Is there a small, dim star in the neighborhood of the Sun that comes close periodically and stirs up the Oort Cloud, sending a much higher than average number of comets into the inner Solar System? Some scientists think the fossil data show that there have been periodic mass extinctions on Earth, and they have investigated astronomical causes of the extinctions. One hypothesis is that a distant companion to the Sun has a wide orbit and periodically has come closer to the Sun as they both have traveled around the Milky Way Galaxy and has stirred up the Oort Cloud as a result. This Oort Cloud disturbance sent a large number of comets to the inner Solar System, which could have caused many impacts on Earth, leading to a mass extinction similar to the one that wiped out the dinosaurs. But some scientists dispute that Earth's extinctions have occurred at regular intervals, and recent surveys with infrared telescopes have found no evidence of a small companion star.

Questions and Problems

Test Your Understanding

1. This chapter deals with leftover planetesimals. What became of most of the others?
   a. They evaporated.
   b. They left the Solar System.
   c. They became part of larger bodies.
   d. They fragmented into smaller pieces.

2. The three types of meteorites come from different parts of their parent bodies. Stony-iron meteorites are rare because
   a. they are hard to find.
   b. the volume of a differentiated body that has both stone and iron is small.
   c. there is very little iron in the Solar System.
   d. the magnetic field of the Sun attracts the iron.

3. As a comet leaves the inner Solar System, the ion tail points
   a. back along the orbit.
   b. forward along the orbit.
   c. toward the Sun.
   d. away from the Sun.

4. Congress tasked NASA with searching for near-Earth objects because
   a. they might impact Earth, as others have in the past.
   b. they are close by and easy to study.
   c. they are moving fast.
   d. they are scientifically interesting.

5. Meteorites can provide information about all of these except
   a. the early composition of the Solar System.
   b. the composition of asteroids.
   c. the composition of comets.
   d. the Oort Cloud.

6. Perihelion is the point in an orbit ___________ the Sun; aphelion is the point in an orbit ___________ the Sun.
   a. closest to; farthest from
   b. farthest from; closer to
   c. at one focus of; at the other focus of

7. Kuiper Belt objects (KBOs) are actually comet nuclei. Why do they not display comae and tails?
   a. Most of the material has already been stripped from the objects.
   b. They are too far from the Sun.
   c. They are too close to the Sun.
   d. The comae and tails are pointing away from Earth, behind the object.

8. Asteroids are small
   a. rock and metal objects orbiting the Sun.
   b. icy objects orbiting the Sun.
   c. rock and metal objects found only between Mars and Jupiter.
   d. icy bodies found only in the outer Solar System.

9. Aside from their periods, short-period and long-period comets differ because
   a. short-period comets orbit prograde, while long-period comets orbit in either sense.
   b. short-period comets contain less ice, while long-period comets contain more.
   c. short-period comets do not develop ion tails, while long-period comets do.
   d. short-period comets come closer to the Sun at closest approach than long-period comets.

10. On average, a bright comet appears about once each decade. Statistically, this means that
    a. one will definitely be observed every tenth year.
    b. one will definitely be observed in each 10-year period.
    c. exactly 10 comets will be observed in a century.
    d. about 10 comets will be observed in a century.

11. Most asteroids are located between the orbits of
    a. Earth and Mars.
    b. Mars and Jupiter.
    c. Jupiter and Saturn.
    d. the Kuiper Belt and the Oort Cloud.

12. Comets, asteroids, and meteoroids may be responsible for delivering a significant fraction of the current supply of ___________ to Earth.
    a. mass
    b. water
    c. oxygen
    d. carbon

13. An iron meteorite most likely came from
    a. an undifferentiated asteroid.
    b. a differentiated asteroid.
    c. a planet.
    d. a comet.

14. Meteor showers occur because Earth passes through the path of
    a. another planet.
    b. a planetesimal.
    c. a comet.
    d. the Moon.

15. Dwarf planets differ from the other planets in that they
    a. have no atmosphere.
    b. have no moons.
    c. are all very far from the Sun.
    d. have lower mass.
    e. are covered in ice.

Thinking about the Concepts

16. Describe ways in which Pluto differs significantly from the Solar System planets.

17. By what criteria did Pluto fail to be considered a planet under the new IAU definition? Explain how this decision demonstrates the self-correcting nature of science.

18. How does the composition of an asteroid differ from that of a comet nucleus?

19. Define *meteoroid*, *meteor*, and *meteorite*.

20. What are the differences between a comet and a meteor in terms of their size, distance, and how long they remain visible?

21. Most meteorites are 4.54 billion years old. Carbonaceous chondrites, however, are 20 million years older. What determines the time of "birth" of these pieces of rock? What does this information tell you about the history of their parent bodies?

22. Most asteroids are found between the orbits of Mars and Jupiter, but astronomers are especially interested in the relative few whose orbits cross that of Earth. Why?

23. How could you and a friend, armed only with your cell phones and knowledge of the night sky, prove conclusively that meteors are an atmospheric phenomenon?

24. Suppose you find a rock that has all the characteristics of a meteorite. You take it to a physicist friend who confirms that it is a meteorite but says that radioisotope dating indicates an age of only a billion years. What might be the origin of this meteorite?

25. Describe differences between the Kuiper Belt and the Oort Cloud as sources of comets. What is the ultimate fate of a comet from each of these reservoirs?

26. What are the three parts of a comet? Which part is the smallest in radius? Which is the most massive?

27. In 1910, Earth passed directly through the tail of Comet Halley. Among the various gases in the tail was hydrogen cyanide, deadly to humans. Yet nobody became ill from this event. Why?

28. Comets have two types of tails. Describe them and explain why they sometimes point in different directions.

29. What is zodiacal light, and what is its source?

30. How might comets and asteroids have contributed to the origin of life on Earth?

Applying the Concepts

31. Comet 67P/Churyumov-Gerasimenko has a diameter of 4 km and a mass of 10^{13} kg.
    a. What is the density of the comet? How does that compare with the density of water?
    b. What is the escape velocity from the surface of this comet?

32. Ceres has a diameter of 975 km and a period of about 9 hours. What is the rotational speed of a point on the surface of this dwarf planet?

33. Figure 12.12 shows the scale of the Solar System out to the Oort Cloud. Judging from this figure, what fraction of the distance between the Sun and Proxima Centauri is occupied by the Oort Cloud, a part of the Solar System?

34. Follow Working It Out 12.1 (and use the information in Appendix 4) to find the perihelion and aphelion distances for Pluto and Eris.

35. Follow Working It Out 12.2 to find the impact energy (in joules) of an asteroid with a mass of 4.6×10^{11} kg traveling at 40 km/s. Does this energy depend on the "target" of impact? What is the equivalent in 1-megaton H bombs?

36. Earth's Moon has a diameter of 3,474 km and orbits at an average distance of 384,400 km. At this distance. it subtends an angle just slightly larger than half a degree in Earth's sky. Pluto's moon Charon has a diameter of 1,186 km and orbits at a distance of 19,600 km from the dwarf planet.
    a. Compare the appearance of Charon in Pluto's skies with the Moon in Earth's skies.
    b. Describe where in the sky Charon would appear as seen from various locations on Pluto.

37. One recent estimate concludes that nearly 800 meteorites with mass greater than 100 grams (massive enough to cause personal injury) strike the surface of Earth each day. Assuming you present a target of 0.25 square meter (m^2) to a falling meteorite, what is the probability that you will be struck by a meteorite during your 100-year lifetime? (Note that the surface area of Earth is approximately 5×10^{14} m^2.)

38. Electra is a 182-km-diameter asteroid accompanied by a small moon orbiting at a distance of 1,350 km in a circular orbit with a period of 3.92 days.
    a. What is the mass of Electra?
    b. What is Electra's density?

39. Calculate the orbital radius of the Kirkwood gap that is in a 3:1 orbital resonance with Jupiter.

40. The orbital periods of Comets Encke, Halley, and Hale-Bopp are 3.3 years, 76 years, and 2,530 years, respectively. Their orbital eccentricities are 0.847, 0.967, and 0.995, respectively.
    a. What are the semimajor axes (in astronomical units) of the orbits of these comets?
    b. What are the minimum and maximum distances from the Sun (in astronomical units) reached by Comets Halley and Hale-Bopp in their respective orbits?
    c. Which region of the Solar System did each likely come from?
    d. Which would you guess is the most pristine comet among the three? Which is the least? Explain your reasoning.

41. Comet Halley has a mass of approximately 2.2×10^{14} kg. It loses about 3×10^{11} kg each time it passes the Sun.
    a. The first confirmed observation of the comet was made in 240 BCE. Assuming a constant period of 76.4 years, how many times has it reappeared since that early sighting?
    b. How much mass has the comet lost since 240 BCE?
    c. What percentage of the comet's total mass today does this amount represent?

42. If Comet Halley is approximated as a sphere 5 km in radius, what is its density if it has a mass of 2.2×10^{14} kg? How does that density compare to that of water (1,000 kg/m^3)?

43. A cubic centimeter of the air you breathe contains about 10^{19} molecules. A cubic centimeter of a comet's tail may typically contain 200 molecules. Calculate the cubic volume of comet tail material that would hold 10^{19} molecules.

44. Some near-Earth objects are in binary systems, so it is possible to estimate their mass. How much energy would be released if a near-Earth asteroid with mass $m = 4.6 \times 10^{11}$ kg hit Earth at a speed (v) of 5 km/s?

45. The estimated amount of zodiacal dust in the Solar System remains constant at approximately 10^{16} kg. Yet zodiacal dust is constantly being swept up by planets or removed by the pressure of sunlight.
    a. If all the dust disappeared (at a constant rate) over a span of 30,000 years, what would the average production rate, in kilograms per second, have to be to maintain the current content?
    b. Is this an example of static or dynamic equilibrium? Explain your answer.

USING THE WEB

46. Dwarf planets:
    a. Go to planetary astronomer Mike Brown's website of dwarf planets (http://gps.caltech.edu/~mbrown/dps.html). How many dwarf planets does he think are in the Solar System? Why is it difficult officially to certify an object as a dwarf planet?
    b. Go to the website for the *New Horizons* mission (http://pluto.jhuapl.edu), which reached Pluto in 2015 and is scheduled to visit Kuiper Belt objects afterward. Click on "Where Is New Horizons?" What is the spacecraft's current location? How far is it from Earth, and how far from Pluto? How long would it take to send a radio signal to the spacecraft? Click on "News Center." What has been learned from this mission?

47. Go to the website for the *Dawn* mission (http://dawn.jpl.nasa.gov).
    a. Read the sections on "Mission" and "Science," and look at the videos and images. What was learned about Vesta on this mission?
    b. What was learned about dwarf planet Ceres during *Dawn*'s visit in 2015?

48. Citizen science projects:
    a. Go to Asteroid Zoo (http://www.asteroidzoo.org/). What are the science goals of this project? Click on "Classify" and read through the Tutorial and Guide. Classify some frames, and save a copy for your homework.
    b. Go to Cosmoquest (https://cosmoquest.org), and click on Asteroid Mappers. What are the science goals? Read through the FAQ and the Tutorial. If you don't already have an account (from Moon Mappers), create one. Log in and get some images, and mark some craters. Do they have data on Ceres? If so, analyze some of those images too.

49. Go to NASA's Asteroid Watch website (http://www.jpl.nasa.gov/asteroidwatch). What is new? Has there been a new discovery or a recent flyby? Was the asteroid studied with a spacecraft, an orbiting telescope, or a ground-based telescope? What has been learned about the object?

50. Go to the Space Weather website (http://spaceweather.com). Are any comets currently visible with the naked eye? Scroll down to "Near Earth Asteroids." Are any "close encounters" coming up in the next few months? Click on a few asteroid names to access the JPL Small-Body Database, where you can view an animation of the orbits. In each case, how close will the NEO be to Earth when it is at its closest? Note the values of e and a in the table under the orbit. Calculate the NEO's closest and farthest distances from the Sun. How large is the object?

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

Astronomers often discover asteroids and other small Solar System objects by comparing two (or more) images of the same star field and looking for bright spots that have moved between the images. The four images in **Figure 12.27** are all "negative images": every dark spot would actually be bright on the sky, and all the white space is dark sky. A negative image sometimes helps the observer pick out faint details, and it is preferable for printing and photocopying. Study these four images. Can you find an asteroid that moves across the field in these images? That's the hard way to do it. A much easier method is to use a "blink comparison," which lets you look at one image and then another very quickly. Make three photocopies of each of these images, cut out each one, and align all of them carefully on top of one another in sequence, so that the stars overlap. You should have 12 pieces of paper, in this order: Image 1, Image 2, Image 3, Image 4; Image 1, Image 2, . . . and so on. Staple the top edge and flip the pages with your thumb, looking carefully at the images. Can you find the asteroid now?

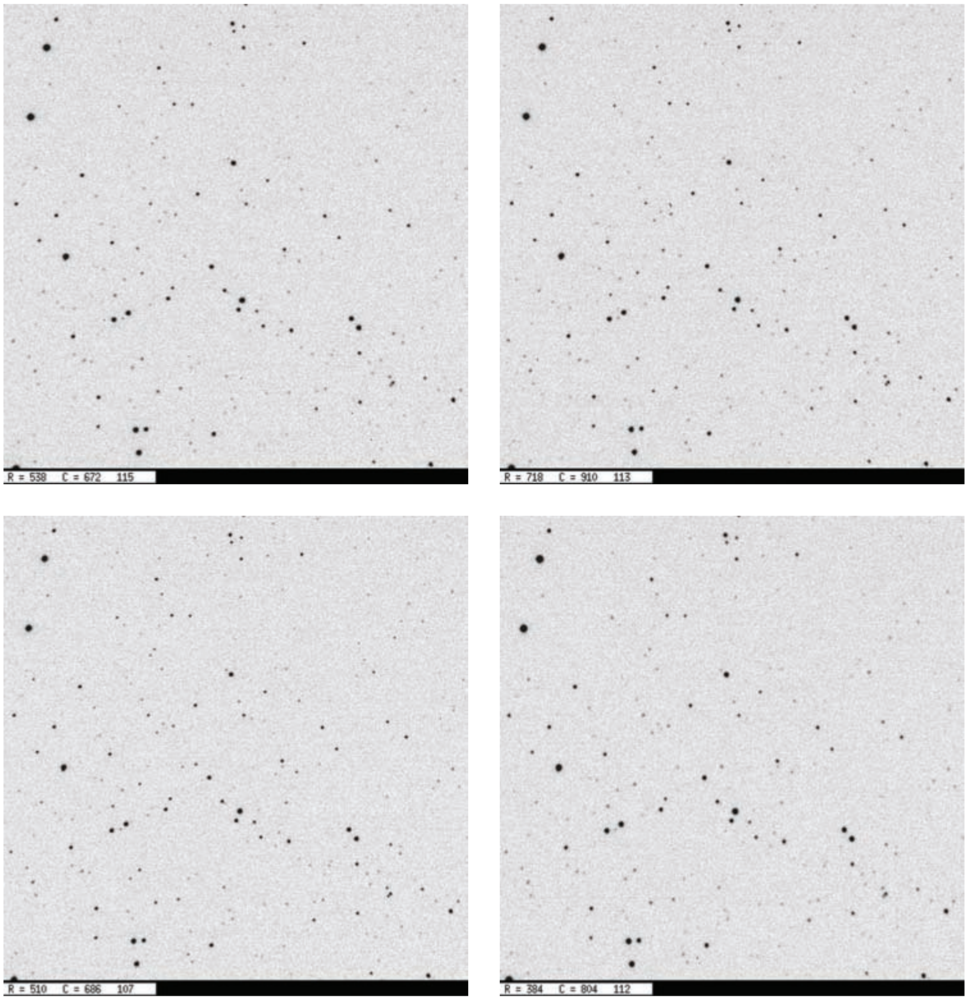


Figure 12.27

1 Circle the asteroid in each image.

The "blink comparison" method takes advantage of a feature of the human eye-brain connection. Humans are much better at noticing things that move than things that do not.

2 Why might this feature be a helpful evolutionary adaptation?

____

____

There's a third method, which sometimes makes things easier to see but requires high-quality digital images. In that method, one image is subtracted from another.

3 If you used the subtraction method with two of these images, what would you expect to see in the resulting image?

____

____

13 Taking the Measure of Stars

To all but the largest telescopes, even nearby stars are just points of light in the night sky. Astronomers study the stars by observing their light, by using the laws of physics discussed in earlier chapters, and by finding patterns in subgroups of stars that are extrapolated to other stars. Astronomers use knowledge of geometry, radiation, and orbits to begin to answer basic questions about stars, such as how they are similar to or different from the Sun, and whether they might have planets orbiting around them as the Sun does.

LEARNING GOALS

By the conclusion of this chapter, you should be able to:

LG 1 Explain how the brightness of nearby stars and their distances from Earth are used to determine how luminous they are.

LG 2 Explain how astronomers obtain the temperatures, sizes, and composition of stars.

LG 3 Describe how astronomers estimate the masses of stars.

LG 4 Classify stars, and organize this information on a Hertzsprung-Russell (H-R) diagram.

LG 5 Explain how the mass and composition of a main-sequence star determine its luminosity, temperature, and size.

The constellation Orion. The reddish object in the middle of the vertical line of stars is a nebula. ▶▶▶

Why do stars have different colors?

13.1 Astronomers Measure the Distance, Brightness, and Luminosity of Stars

When looking up at the stars in the sky, it is immediately noticeable that they differ in brightness and color. However, we don't know if one star appears brighter than another because it has a higher luminosity and is emitting more light or because it is closer to us. In this section, you will learn how to find the distances to nearby stars and how to use distance and apparent brightness to find the luminosity of a star.

Stereoscopic Vision

Your two eyes have different views that depend on the distance to the object you are viewing. Hold up your finger in front of you, quite close to your nose. View it with your right eye only and then with your left eye only. Each eye sends a slightly different image to your brain, so your finger *appears* to move back and forth relative to the background behind it. Now hold up your finger at arm's length, and blink your right eye, then your left. Your finger appears to move much less. The way your brain combines the different information from your eyes to perceive the distances to objects around you is called **stereoscopic vision**. **Figure 13.1a** shows an overhead view of the experiment you just performed with your finger. The left eye sees the blue pencil almost directly between the green balls on the bookshelf. But the right eye sees the blue pencil to the left of both balls. Similarly, the position of the pink pencil appears to vary. Because the pink pencil is closer to the observer, its position appears to change more than the position of the blue pencil—it seems to move from the right of the blue pencil to the left of the blue pencil.

Stereoscopic vision enables you to judge the distances of objects as far away as a few hundred meters, but beyond that it is of little use. Your right eye's view of a mountain several kilometers away is indistinguishable from the view seen by your left eye—all you can determine is that the mountain is too far away for you to judge its distance stereoscopically. The distance over which your stereoscopic vision works is limited by the separation between your eyes, about 6 centimeters (cm). If you could separate your eyes by several meters, the view from each eye would be different enough for you to judge the distances to objects that are kilometers away.

Astronomy in Action: Parallax

Of course, you cannot separate your eyes, but you can compare pictures taken with a camera from two widely separated locations. The greatest separation we can obtain without leaving Earth is to let Earth's orbital motion carry us from one side of the Sun to the other. If you take a picture of the sky tonight and then wait 6 months and take another picture, the distance between the two locations is the diameter of Earth's orbit (2 astronomical units [AU]), which gives us more powerful stereoscopic vision.

Distances to Nearby Stars

Figure 13.1b shows how astronomers apply this concept of stereoscopic vision to measure the distances to stars. This illustration shows Earth's orbit as viewed from far above the Solar System. The change in position of Earth over 6 months is like the distance between the right eye and the left eye in Figure 13.1a. The nearby (pink and blue) stars are like the pink and blue pencils, while the distant yellow stars are like the green balls on the bookcase. Because of the shift in perspective as Earth orbits the Sun, nearby stars appear to shift their positions. The pink star, which is closer, appears to move farther than the more distant blue star. Over the

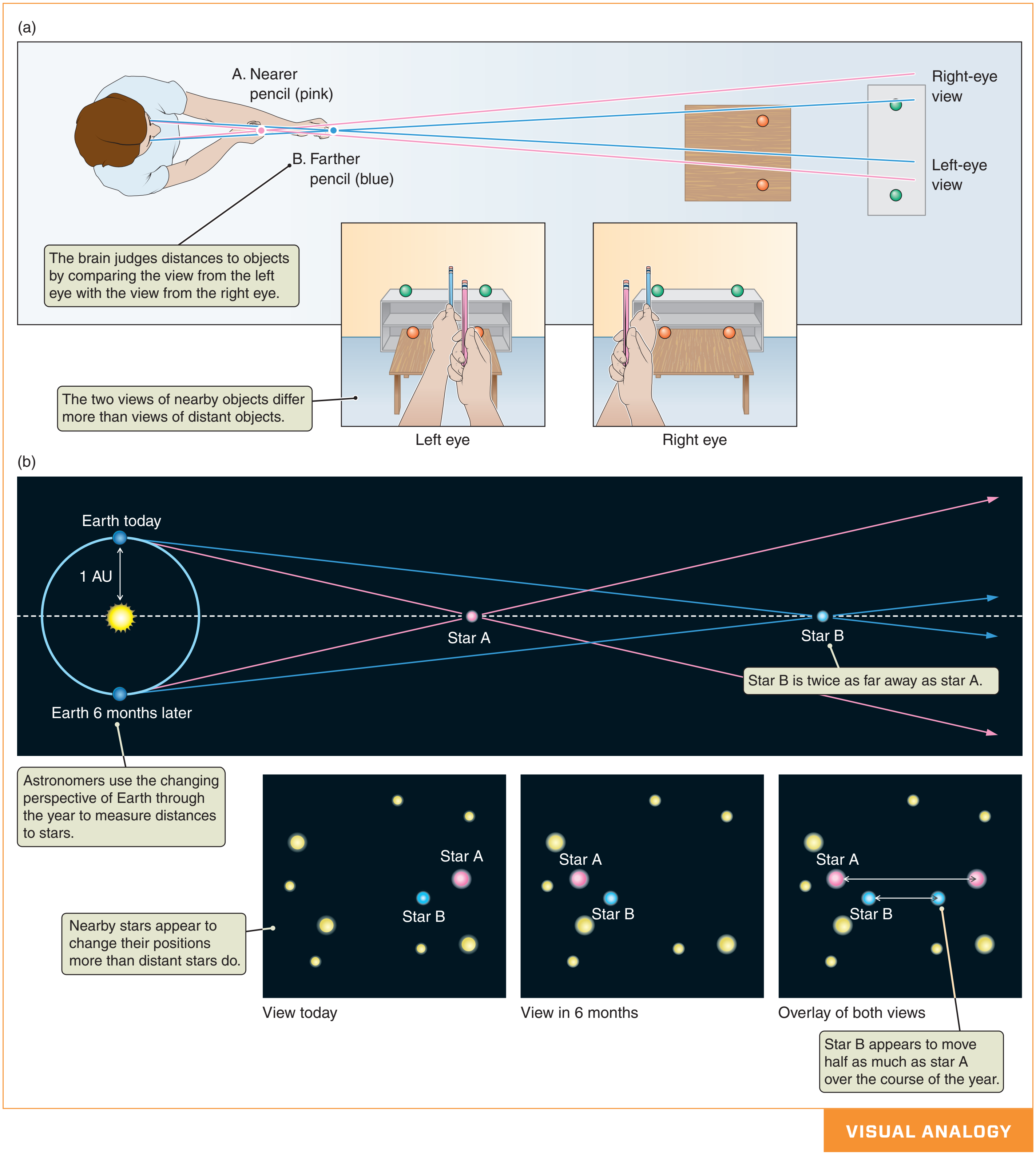


Figure 13.1 (a) Stereoscopic vision enables you to determine the distance to an object by comparing the view from each eye. (b) Similarly, comparing views from different places in Earth's orbit enables astronomers to determine the distances to stars. As Earth moves around the Sun, the apparent positions of nearby stars change more than the apparent positions of more distant stars. (The diagram is not to scale.)

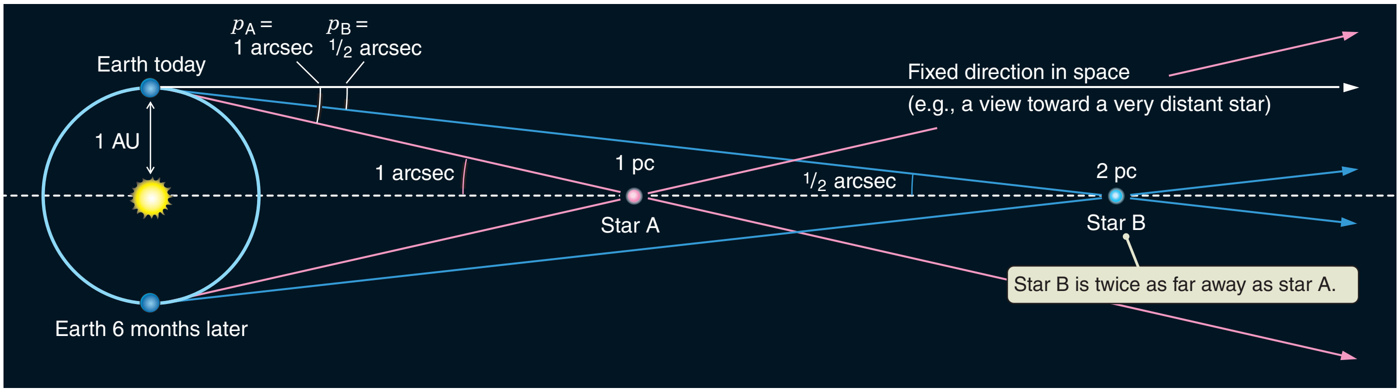


Figure 13.2 The parallax (p) of a star is inversely proportional to its distance. More distant stars have smaller parallaxes. (The diagram is not to scale.)

Nebraska Simulation: Parallax Calculator

course of 1 full year, the nearby star appears to move one way and then back again with respect to distant background stars, returning to its original position at the end of that 1-year period. We can determine the distance to the star using the amount of this apparent shift and geometry.

The eye cannot detect the changes in position of a nearby star throughout the year, but telescopes can reveal these small shifts relative to the background stars. **Figure 13.2** shows Earth, the Sun, and two stars in a similar configuration as that in Figure 13.1b. Look first at star A, the closest star. When Earth is at the top of the figure, it forms a right triangle with the Sun and star A at the other corners. (Remember that a right triangle is one with a 90° angle in it.) The short leg of the triangle is the distance from Earth to the Sun, which is 1 AU. The long leg of the triangle is the distance from the Sun to star A. The small angle near star A is called the *parallactic angle*, or simply **parallax**, of the star. As Earth completes an orbit around the Sun, the star's position in the sky appears to shift back and forth, returning to its original position at the end of that year. The amount of this shift is equal to twice the parallax.

More distant stars make longer and skinnier triangles with smaller parallaxes. Star B is twice as far away as star A, and its parallax is only half the parallax of star A. If you were to draw a number of such triangles for different stars, you would find that increasing the distance to the star always reduces the star's parallax. Moving a star 3 times farther away reduces its parallax to $\frac{1}{3}$ of its original value. Moving a star 10 times farther away reduces its parallax to $\frac{1}{10}$ of its original value. The parallax of a star (p) is inversely proportional to its distance (d).

The parallaxes of real stars are tiny. Recall from Chapter 2 that the full circle of the sky can be divided into 360 degrees. The apparent diameter of the full Moon in the sky averages about half a degree. Just as an hour on the clock is divided into minutes and seconds, a degree of sky can be divided into arcminutes and arcseconds. An **arcminute** (abbreviated **arcmin**) is 1/60 of a degree, and an **arcsecond** (abbreviated **arcsec**) is 1/60 of an arcminute. An arcsecond is about equal to the angle formed by the diameter of a golf ball at a distance of 9 km.

Recall from Chapter 1 that we usually use units of light-years to indicate distances to stars. One light-year is the distance that light travels in 1 year—about 9.5 trillion kilometers (km). We use this unit because it is the unit you are most likely to see online or in a popular book about astronomy. When astronomers discuss distances to stars and galaxies, however, the unit they often use is the **parsec (pc)**, which is equal to 3.26 light-years (or 206,265 AU). The term is short for *parallax second*—a star at a distance of 1 parsec has a parallax of 1 arcsecond.

13.1 Working It Out Parallax and Distance

As seen in Figure 13.2, if star B is twice as far away from us as star A, then star B will have half the parallax of star A. The parallax of a star (p) is inversely proportional to its distance (d):

$$p \propto \frac{1}{d} \quad \text{or} \quad d \propto \frac{1}{p}$$

If the angle at the apex of a triangle is 1 arcsec, and the base of the triangle is 1 AU, then the length of the triangle is 1 parsec. One parsec equals 206,265 AU (see Appendix 1), or 3.09×10^{16} meters, or 3.26 light-years. Astronomers use the unit of parsecs because it makes the relationship between distance and parallax easier than that with the use of light-years.

As illustrated in Figure 13.2, a star with a parallax of 1 arcsec is at a distance of 1 pc. The inverse proportionality between distance and parallax becomes

$$\begin{pmatrix}\text{Distance measured}\\ \text{in parsecs}\end{pmatrix} = \frac{1}{\begin{pmatrix}\text{Parallax measured}\\ \text{in arcseconds}\end{pmatrix}}$$

or

$$d(\text{pc}) = \frac{1}{p(\text{arcsec})}$$

Suppose that the parallax of a star is measured to be 0.5 arcsec. The distance can be found by

$$d(\text{pc}) = \frac{1}{0.5} = 2\,\text{pc}$$

Similarly, a star with a measured parallax of 0.01 arcsec is located at a distance of $1/0.01 = 100$ pc.

After the Sun, the next closest star to Earth is Proxima Centauri. Located at a distance of 4.24 light-years, Proxima Centauri is a faint member of a system of three stars called Alpha Centauri. What is this star's parallax? First, we convert the distance to parsecs:

$$d = 4.24\,\text{light-years} \times \frac{1\,\text{parsec}}{3.26\,\text{light-years}} = 1.30\,\text{parsecs}$$

Then,

$$p(\text{arcsec}) = \frac{1}{1.30\,\text{pc}} = 0.77\,\text{arcsec}$$

Even the closest star to the Sun has a parallax of only about $\frac{3}{4}$ arcsec.

When astronomers began to measure the parallax angles of stars, they discovered that stars are very distant objects (**Working It Out 13.1**). The first successful measurement of the parallax of a star was made by F. W. Bessel (1784–1846), who in 1838 reported a parallax of 0.314 arcsec for the star 61 Cygni. This finding implied that 61 Cygni was 3.2 pc away, or 660,000 times as far away as the Sun. With this one measurement, Bessel increased the known volume of the universe by a factor of 10,000. Today, astronomers know of about 60 stars in 54 single-, double-, or triple-star systems within 5 pc (16.3 light-years) of the Sun. In the neighborhood of the Sun, each star or star system has on average a volume of about 260 cubic light-years of space to itself.

Most stars are so far away that the parallax angle is too small to measure using ground-based telescopes, which are limited by Earth's atmosphere. In the 1990s, the European Space Agency's Hipparcos satellite measured the positions and parallaxes of 120,000 stars, thus greatly improving our picture of the Sun's stellar neighborhood. The accuracy of any given Hipparcos parallax measurement is about ± 0.001 arcsec. Because of this observational uncertainty, measurements of the distances to stars are not perfect. For example, a star with a Hipparcos-measured parallax of 0.004 ± 0.001 arcsec really has a parallax between 0.003 and 0.005 arcsec. This gives a corresponding distance range of 200–330 pc from Earth. As an analogy, consider your speed while driving down the road. If your digital speedometer says 10 kilometers per hour (km/h), you might actually be traveling 10.4 km/h or 9.6 km/h. The precision of your speedometer is limited to the nearest 1 km/h.

A successor to Hipparcos is Gaia, a space mission to study stellar parallaxes that was launched at the end of 2013. Gaia is expected to observe 1 billion stars and measure the parallaxes of 20 million of them with a high precision. Other methods of measuring distance to more remote stars will be discussed later.

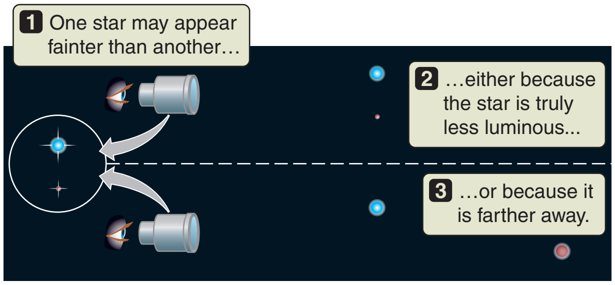


Figure 13.3 The brightness of a star visible in our sky depends on both its luminosity—how much light it emits—and its distance.

Luminosity, Brightness, and Distance

The stars in Earth's sky are of different brightnesses. In Chapter 5, you saw that brightness corresponds to the amount of energy falling on a square meter of area each second in the form of electromagnetic radiation. Although the brightness of a star can be measured directly, it does not immediately give much information about the star itself. As illustrated in **Figure 13.3**, a bright star in the night sky may in fact only appear bright because it is nearby. Conversely, a faint star may be a powerful beacon, still visible despite its tremendous distance.

Astronomers measure the brightness of stars by comparing them to one another. The system they use dates back 2,100 years, when the Greek astronomer Hipparchus classified stars according to their brightness. The details of his system, which is still in use today, are discussed in **Working It Out 13.2** and Appendix 7.

To learn about the actual properties of a star, astronomers need to know the total energy radiated by a star each second—the star's luminosity. Recall from

13.2 Working It Out The Magnitude System

The **magnitude** system of brightness for celestial objects can be traced back 2,100 years to the ancient Greek astronomer Hipparchus, who classified the brightest stars he could see as being "of the first magnitude" and the faintest as being "of the sixth magnitude." Later, astronomers defined Hipparchus's 1st magnitude stars as being exactly 100 times brighter than his 6th magnitude stars. Hipparchus must have had typical eyesight, because an average person under dark skies can see stars only as faint as 6th magnitude. Today, telescopes extend our vision far into space. The Hubble Space Telescope can integrate for long exposures and detect stars as faint as 30th magnitude. The limits have also been extended to stars brighter than 1st magnitude using zero and negative numbers. A negative magnitude signifies that an object is *brighter* than an object at zero magnitude. For example, Sirius, the brightest star in the sky, has a magnitude of −1.46. Venus can be as bright as magnitude −4.4, or about 15 times brighter than Sirius and bright enough to cast a shadow. The magnitude of the full Moon is −12.6, and that of the Sun is −26.7 (**Figure 13.4**).

Looking at this mathematically, with five steps between the 1st and 6th magnitudes, each step is equal to the fifth root of 100, or $100^{1/5}$, which is approximately 2.512. This system is logarithmic, but instead of the usual base 10, it is base 2.512. Thus, a 5-magnitude difference in brightness equals $(2.512)^5$, or 100, times difference in brightness. Fifth magnitude stars are 2.512 times brighter than 6th magnitude stars, and 4th magnitude stars are $2.512 \times 2.512 = 6.310$ times brighter than 6th magnitude stars. A 2.5-magnitude difference equals $(2.512)^{2.5}$, or 10, times difference in brightness. The brightness ratio between any two stars is equal to $(2.512)^N$, where N is the magnitude difference between them.

Since the limit of the Hubble Space Telescope (HST) is 30 and $30 - 6 = 24$, HST can detect stars that are $(2.512)^{24} = 4 \times 10^9$, or 4 billion, times fainter than the magnitude 6 that the naked eye can see. Or if we compare the Sun and the Moon, the Sun is 14 magnitudes, or $(2.512)^{14} = 4 \times 10^5$ (400,000) times brighter than the full Moon. (More detailed calculations and a table of magnitudes and brightness differences are located in Appendix 7.)

The magnitude of a star, as we have discussed it, is called the star's **apparent magnitude** because it is the brightness of the star as it *appears* in Earth's sky. Now imagine that all stars were located at exactly 10 pc (32.6 light-years) away from Earth. The brightness of each star would then reflect its luminosity. If the distance from Earth to a star is known, astronomers compute how bright the star would appear if it were located at 10 pc. The **absolute magnitude** of a star—its apparent magnitude at a distance of 10 pc—measures the star's luminosity.

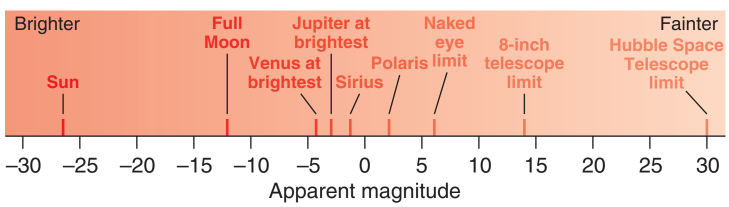


Figure 13.4 Apparent magnitude indicates the apparent brightness of an object in our sky. The brightest objects have a negative apparent magnitude, while telescopes have extended the observable range to fainter objects with higher magnitudes.

Chapter 5 that the brightness of an object that has a known luminosity and is located at a distance d is given by the following equation:

$$\text{Brightness} = \frac{\text{Total light emitted per second}}{\text{Area of a sphere of radius } d} = \frac{\text{Luminosity}}{4\pi d^2}$$

You can rearrange this equation, moving the quantities you know how to measure (distance and brightness) to the right-hand side and the quantity you would like to know (luminosity) to the left, to get

$$\text{Luminosity} = 4\pi d^2 \times \text{Brightness}$$

This equation is used to find how much total light a star must be giving off in order to appear as bright as it does when seen from Earth.

Different stars have different luminosities. The Sun provides a convenient comparison when measuring the properties of stars, including their luminosity. The luminosity of the Sun is measured at $L_{\text{Sun}} = 3.9 \times 10^{26}$ watts (W). The most luminous stars exceed a million times the luminosity of the Sun ($10^6\ L_{\text{Sun}}$). The least luminous stars have luminosities less than 1/10,000 that of the Sun ($10^{-4}\ L_{\text{Sun}}$). The most luminous stars are therefore more than 10 billion (10^{10}) times more luminous than the least luminous stars. Only a very small fraction of stars are near the upper end of this range of luminosities. The vast majority of stars are at the faint end of this distribution, less luminous than the Sun. **Figure 13.5** shows the relative number of stars compared to their luminosities in solar units. (Distances for the nearest stars are obtained from their parallaxes; other methods—to be discussed later—are used for the more distant stars.)

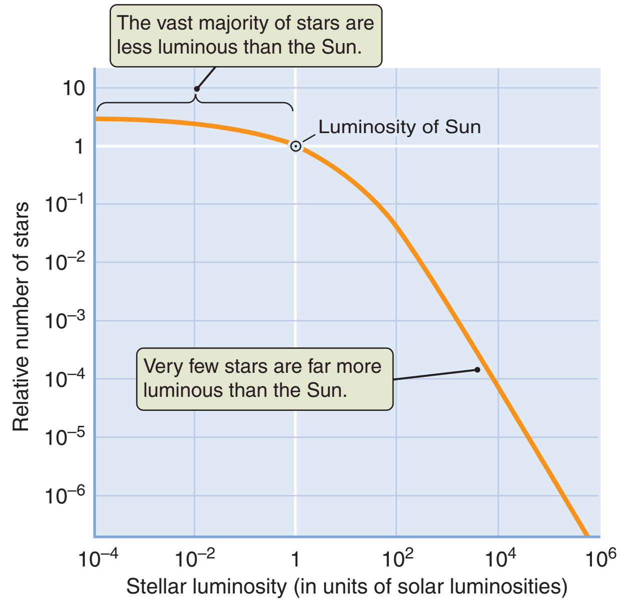


Figure 13.5 The distribution of the luminosities of stars is plotted logarithmically in this graph, so that increments are in powers of 10.

CHECK YOUR UNDERSTANDING 13.1

Stars A and B appear equally bright, but star A is twice as far away from us as star B. Which of the following is true? (a) Star A is twice as luminous as star B. (b) Star A is 4 times as luminous as star B. (c) Star B is twice as luminous as star A. (d) Star B is 4 times as luminous as star B. (e) Star A and star B have the same luminosity because they have the same brightness.

Nebraska Simulation: Stellar Luminosity Calculator

13.2 Astronomers Can Determine the Temperature, Size, and Composition of Stars

Two everyday concepts—stereoscopic vision and the fact that objects appear brighter when closer—have provided the tools needed to measure the distance and luminosity of the closest stars. Stars that appear to be faint points of light in the night sky are in fact luminous beacons located at great distances. The laws of radiation that we described in Chapter 5 reveal still more about stars.

Stars are gaseous, but they are dense enough that the radiation from a star comes close to obeying the same laws as the radiation from solid objects like the heating element on an electric stove. We can therefore use our knowledge of Planck blackbody radiation to understand the radiation from stars. Recall both the Stefan-Boltzmann law from Chapter 5, which states that among same-sized objects, the hotter objects are more luminous, and Wien's law, which states that hotter objects are bluer. In this section, we will use these two laws to measure the temperatures and sizes of stars. We will also develop a more detailed understanding of the line emission mentioned in Chapter 5 to obtain information about the composition of stars.

Wien's Law Revisited: The Color and Surface Temperature of Stars

Wien's law (see Working It Out 5.3) shows that the temperature of an object determines the peak wavelength of its spectrum. The hotter the surface of an object, the bluer the light that it emits. Stars with especially hot surfaces are blue, stars with especially cool surfaces are red, and yellow-white stars such as the Sun are in-between. If you obtain a spectrum of a star and measure the wavelength at which the spectrum peaks, then Wien's law will tell you the temperature of the star's surface. The color of a star tells you about the temperature only at the surface, because this layer is giving off most of the radiation that we see. (Stellar interiors are far hotter than this, as we will discuss in the next chapter.)

▶II **AstroTour:** Stellar Spectrum

In practice, it is not necessary to obtain a complete spectrum of a star to determine its temperature. Astronomers often measure the colors of stars by comparing the brightness at two different, specific wavelengths. The brightness of a star is usually measured through an optical **filter**—sometimes just a piece of colored glass—that lets through only a small range of wavelengths. Two of the most common filters are a blue filter that allows light with wavelengths of about 440 nanometers (nm) to pass through and a "visual" (yellow-green) filter that allows light with wavelengths of about 550 nm to pass through. The ratio of brightness between the blue and visual filters is called the *color index* of the star (more details are discussed in Appendix 7). From a pair of pictures of a group of stars, each taken through a different filter, we can find an approximate value of the surface temperature of every star in the picture—perhaps hundreds or even thousands—all at once. This type of analysis shows that there are many more cool stars than hot stars, and most stars have surface temperatures lower than that of the Sun.

Classification of Stars by Surface Temperature

Although the hot "surface" of a star emits radiation with a spectrum very close to a smooth Planck blackbody curve, this light must then escape through the outer layers of the star's atmosphere. The atoms and molecules in the cooler layers of the star's atmosphere leave their absorption line fingerprints in the escaping light, as shown in **Figure 13.6**. Under some circumstances, the atoms and molecules in the star's atmosphere, along with any gas that might be found near the star, can produce emission lines in stellar spectra. Absorption and emission lines complicate how astronomers use the laws of Planck blackbody radiation to interpret light from stars, but spectral lines provide a wealth of information about the state of the gas in a star's atmosphere.

The spectra of stars were first classified during the late 1800s, long before stars, atoms, or radiation were well understood. Stars were classified by the appearance of the dark bands (now known as absorption lines) seen in their spectra. The original ordering of this classification was arbitrarily based on the prominence of particular absorption lines known to be associated with the element hydrogen. Stars with the strongest hydrogen lines were denoted *A stars*, stars with somewhat weaker hydrogen lines were denoted *B stars*, and so on.

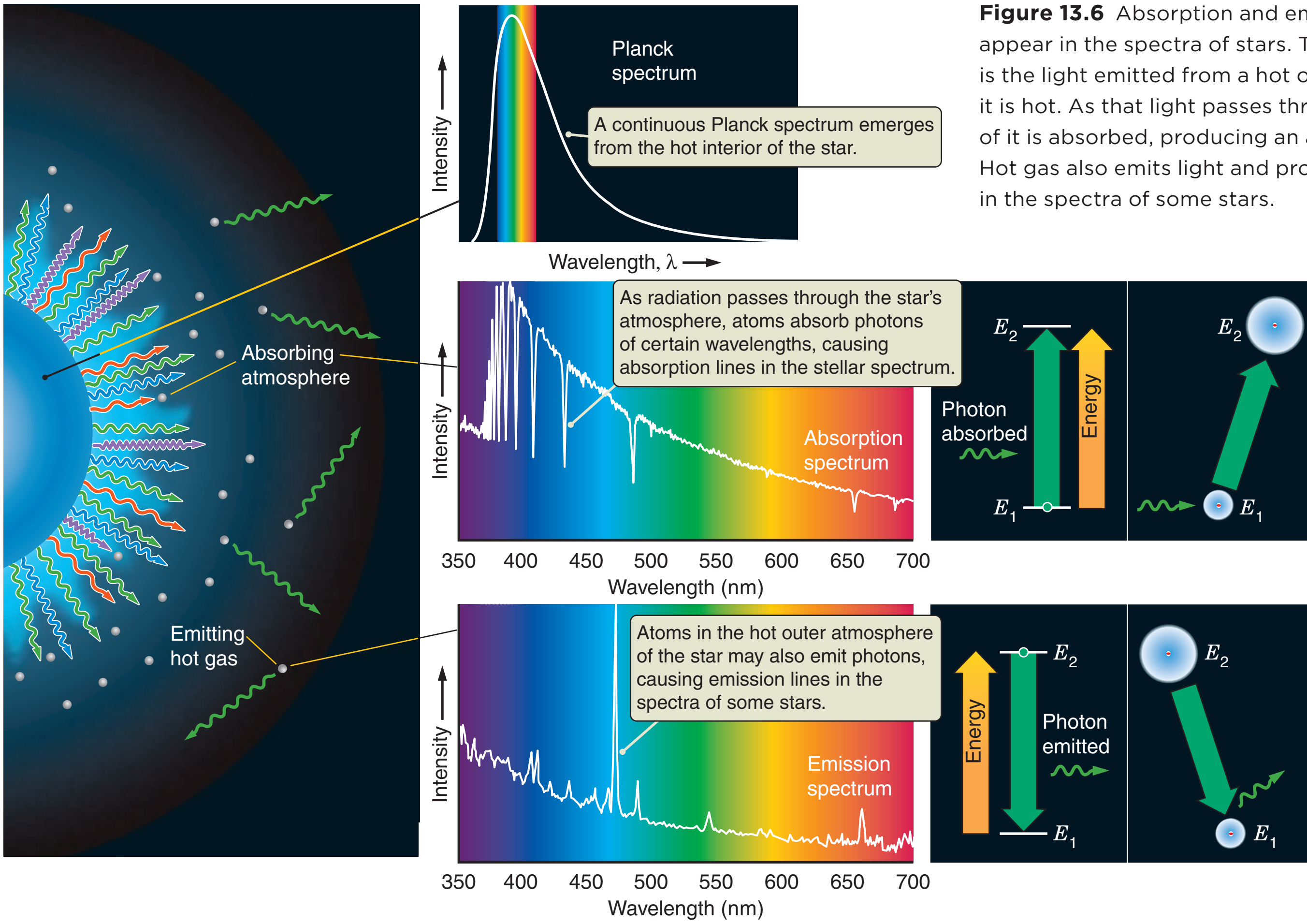


Figure 13.6 Absorption and emission lines both appear in the spectra of stars. The Planck spectrum is the light emitted from a hot object, just because it is hot. As that light passes through a gas, some of it is absorbed, producing an absorption spectrum. Hot gas also emits light and produces emission lines in the spectra of some stars.

Annie Jump Cannon (1863–1941) led an effort at the Harvard College Observatory to examine and classify the spectra of hundreds of thousands of stars systematically. She dropped many of the earlier spectral types, keeping only seven that were subsequently reordered based on surface temperatures. Spectra of stars of different spectral types are shown between the horizontal bars in **Figure 13.7**. The hottest stars, with surface temperatures above 30,000 K, are denoted *O stars*. O stars have only weak absorption lines from hydrogen and helium. The coolest stars—*M stars*—have temperatures as low as about 2800 K. M stars show absorption lines from many different types of atoms and molecules. The sequence of **spectral types** of stars, from hottest to coolest, is O, B, A, F, G, K, M. This sequence has undergone several modifications over time, most recently to add cooler objects known as brown dwarfs with spectral types L, T, and Y.

Astronomers divide the main spectral types into a finer sequence of subclasses by adding numbers to

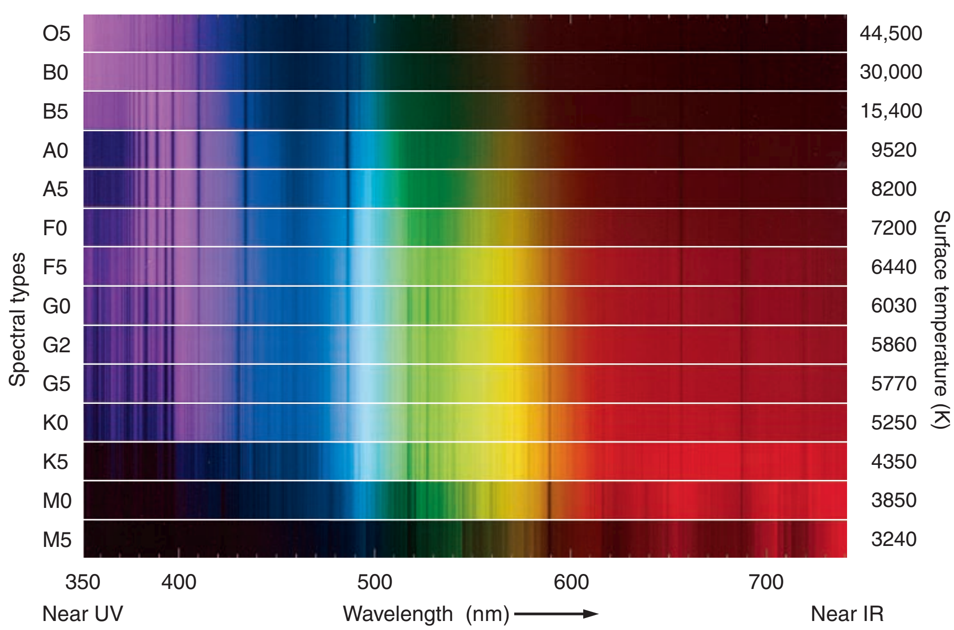


Figure 13.7 Spectra of stars with different spectral types, ranging from hot, blue O stars to cool, red M stars. Hotter stars are more luminous at shorter wavelengths. The dark lines are absorption lines.

the letter designations. For example, the hottest B stars are B0 stars, slightly cooler B stars are B1 stars, and so on. The coolest B stars are B9 stars, which are only slightly hotter than A0 stars. The Sun is a G2 star. The boundaries between spectral types are not always easy to determine. A hotter-than-average G star is very similar to a cooler-than-average F star.

In Figure 13.7, notice that not only are hot stars bluer than cool stars but also the absorption lines in their spectra are quite different. The temperature of the gas in the atmosphere of a star affects the state of the atoms in that gas, which in turn affects the energy level transitions available to absorb radiation (see Section 5.2 to review the concept of atomic energy levels). In O stars, the temperature is so high that most atoms have had one or more electrons stripped from them by energetic collisions within the gas. Few transitions are available in the visible part of the electromagnetic spectrum, so the visible spectrum of an O star is relatively featureless. At lower temperatures, there are more atoms in energy states that can absorb light in the visible part of the spectrum, so the visible spectra of cooler stars are far more complex than the spectra of O stars.

All absorption lines have a temperature at which they are strongest. For example, absorption lines from hydrogen are most prominent at temperatures of about 10,000 K, which is the surface temperature of an A star. At the very lowest stellar temperatures, atoms in the atmosphere of a star react with each other, forming molecules. Molecules such as titanium oxide (TiO) are responsible for much of the absorption in the atmospheres of cool M stars.

Because different spectral lines are formed at different temperatures, astronomers can use these absorption lines to measure a star's temperature. The surface temperatures of stars measured in this way agree extremely well with the surface temperatures of stars measured using Wien's law, again confirming that the physical laws that apply on Earth apply to stars as well.

The Composition of Stars

Most of the variations in the lines of a particular chemical element seen in stellar spectra are due to temperature, but the details of the absorption and emission lines found in starlight also carry a wealth of other information. By applying the physics of atoms and molecules to the study of stellar absorption lines, astronomers can accurately determine not only surface temperatures of stars but also pressures, chemical compositions, magnetic-field strengths, and other physical properties. In addition, by making use of the Doppler shift of emission and absorption lines, astronomers can measure rotation rates, atmospheric motions, expansion and contraction, "winds" driven away from stars, and other dynamic properties of stars.

The chemical composition of a star is also found from its spectra. In Chapter 5, you saw that because each type of atom has different energy levels, each type of atom has different spectral lines. The patterns of spectral lines are measured in laboratories on Earth and then used to identify the types of atoms (or molecules) in stars. For example, if a star has absorption lines that correspond to the energy difference between two levels in the calcium atom, then we know that calcium is present in the atmosphere of the star.

The strengths of various absorption lines tell us not only what kinds of atoms are present in the gas but also the amount of each. However, we must take great care in interpreting spectra to account properly for the temperature and density of the gas in the atmosphere of a star. Recall the "astronomer's periodic table of the elements" from Figure 5.16. Typically, hydrogen composes more than 90 percent of the atoms in the atmosphere of a star, while helium accounts for most of what remains. All of the other chemical elements, which collectively are called **heavy elements** or **massive elements**, are present in only very small amounts.

Table 13.1 shows the chemical composition of the atmosphere of the Sun. The Sun's composition is fairly typical for stars in its vicinity, but the percentages of various heavy elements can vary tremendously from star to star. Some stars have lower amounts of heavier elements than the Sun. The existence of such stars, all but devoid of more massive elements, provides important clues about the origin of chemical elements and the chemical evolution of the universe. Note that many of the atoms that make up Earth and its atmosphere (for example, iron, silicon, nitrogen, oxygen, and carbon) exist as only a small percentage of the Sun.

TABLE 13.1 The Relative Amounts of Different Chemical Elements in the Atmosphere of the Sun

Element	Percentage of Atoms in the Sun This Element Represents	Percentage of Sun's Mass This Element Represents
Hydrogen	92.5	74.5
Helium	7.4	23.7
Oxygen	0.064	0.82
Carbon	0.039	0.37
Neon	0.012	0.19
Nitrogen	0.008	0.09
Silicon	0.004	0.09
Magnesium	0.003	0.06
Iron	0.003	0.16
Sulfur	0.001	0.04
Total of others	0.001	0.03

The Stefan-Boltzmann Law and Finding the Sizes of Stars

Stars are so far away that only two can be imaged as more than point sources of light. To determine the size of a star, astronomers must use other measurements: the temperature and the luminosity.

The temperature of a star can be found directly, either from its color through Wien's law (**Figure 13.8a**) or from the strength of its spectral lines. The temperature of the surface of a star is one factor that influences its luminosity. If a large star and a small star are the same temperature, they will emit the same energy from every patch of surface, but the large star has more patches, so it is more luminous altogether. Conversely, if two stars are the same size, the hotter one will be more luminous than the cooler one. This is an application of the Stefan-Boltzmann law, shown in Figure 13.8b. A small, hot star might even be more luminous than a larger cool star.

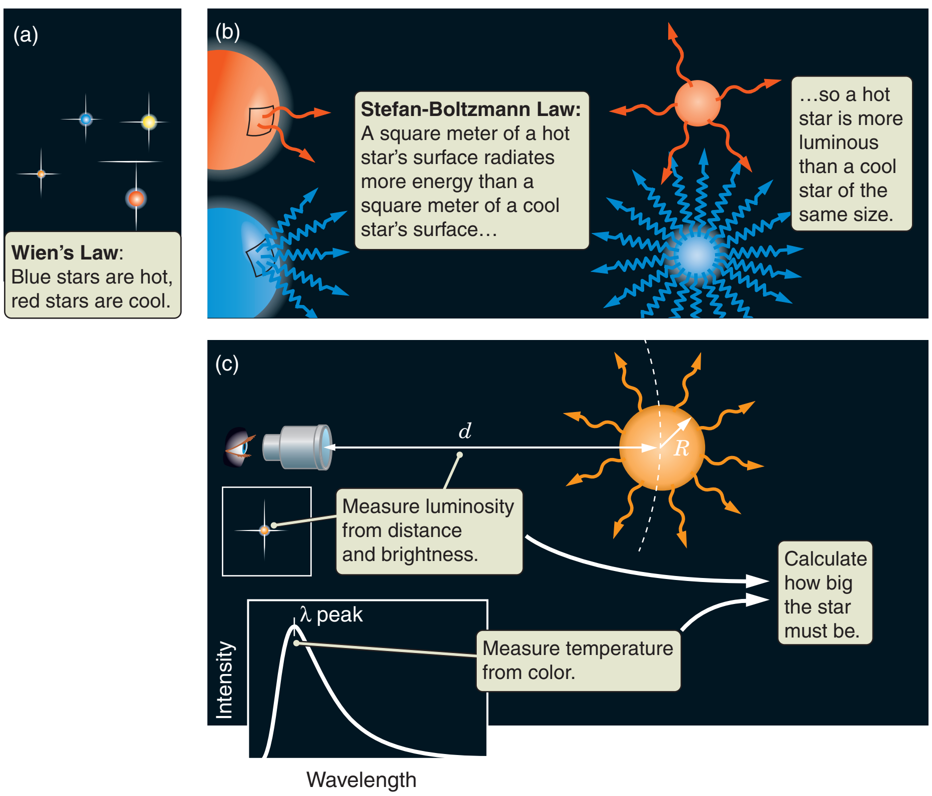


Figure 13.8 (a) The temperature of a star can be found from its color through Wien's law. (b) The luminosity depends on both the temperature and the size of the star. (c) Once the temperature and the luminosity are known, the size of the star can be calculated.

The luminosity of a star can be found from its brightness and its distance. Because the luminosity depends on both the temperature and the size of the star, we can use the luminosity and the temperature to determine the radius of the star, as shown in Figure 13.8c (**Working It Out 13.3**). Astronomers have used the luminosity-temperature-radius relationship to estimate the radii of many thousands of stars. The radius of the Sun, written R_{Sun}, is 696,000 km. One of the

13.3 Working It Out Estimating the Sizes of Stars

In Chapter 5, you learned that according to the Stefan-Boltzmann law, the amount of energy radiated each second by each square meter of the surface of a star is equal to the constant σ multiplied by the surface temperature of the star raised to the fourth power. Written as an equation, this relationship says:

$$\left(\begin{array}{c}\text{Energy radiated each}\\ \text{second by 1 m}^2\text{ of surface}\end{array}\right) = \sigma T^4$$

To find the total amount of light radiated each second by the star, we need to multiply the radiation per second from each square meter by the number of square meters of the star's surface:

$$\left(\begin{array}{c}\text{Energy radiated}\\ \text{each second}\end{array}\right) = \left(\begin{array}{c}\text{Energy radiated each}\\ \text{second by 1 m}^2\text{ of surface}\end{array}\right) \times \left(\begin{array}{c}\text{Surface}\\ \text{area}\end{array}\right)$$

The left-hand term in this equation—the total energy emitted by the star per second (in units of joules per second [J/s] = watts [W])—is the star's luminosity, L. The middle term—the energy radiated by each square meter of the star per second (in units of joules per square meter per second; J/m^2/s)—can be replaced with the σT^4 factor from the Stefan-Boltzmann law. The remaining term—the number of square meters covering the surface of the star—is the surface area of a sphere, $A_{\text{sphere}} = 4\pi R^2$ (in units of square meters; m^2), where R is the radius of the star.

If we replace the words in the equation with the appropriate mathematical expressions for the Stefan-Boltzmann law and the area of a sphere, our equation for the luminosity of a star looks like this:

$$\text{Luminosity} = \sigma T^4 \times 4\pi R^2$$

Combining gives

$$L = 4\pi R^2 \sigma T^4 \text{ J/s (W)}$$

This last equation is called the **luminosity-temperature-radius relationship** for stars. Because the constants (4, π, and σ) do not change, the luminosity of a star is proportional only to R^2T^4. Make a star 3 times as large, and its surface area becomes $3^2 = 9$ times as large. There is 9 times as much area to radiate, so there is 9 times as much radiation. Make a star twice as hot, and each square meter of the star's surface radiates $2^4 = 16$ times as much energy. Larger, hotter stars are more luminous than smaller, cooler stars.

Now turn this question around and ask, How large must a star of a given temperature be to have a total luminosity of L? The star's luminosity (L) and temperature (T) are quantities that we can measure, and the star's radius (R) is what we want to know. We can rearrange the previous equation, moving the properties that we know how to measure (temperature and luminosity) to the right-hand side of the equation and the property that we would like to know (the radius of the star) to the left-hand side. After a couple of steps of algebra, we find:

$$R = \sqrt{\frac{L}{4\pi\sigma T^4}} = \frac{1}{T^2}\sqrt{\frac{L}{4\pi\sigma}}$$

Again, the right-hand side of the equation contains only things that we know or can measure. The constants 4, π, and σ are always the same. We can find L, the luminosity of the star, by use of the measurements of the star's brightness and parallax (although only for nearby stars with known parallax). T is the surface temperature of the star, which can be measured from its color. From the relationship of these measurements, we now know something new: the size of the star.

Often, we compare two stars and the constants all cancel out, leaving L, T, and R:

$$\frac{L_{\text{star 1}}}{L_{\text{star 2}}} = \frac{R^2_{\text{star 1}}}{R^2_{\text{star 2}}} \times \frac{T^4_{\text{star 1}}}{T^4_{\text{star 2}}} \quad \text{or} \quad \frac{R_{\text{star 1}}}{R_{\text{star 2}}} = \sqrt{\frac{L_{\text{star 1}}}{L_{\text{star 2}}}} \times \frac{T^2_{\text{star 2}}}{T^2_{\text{star 1}}}$$

Suppose we compare the Sun to the second brightest star in the constellation Orion, a red star called Betelgeuse. From its spectrum, we know that Betelgeuse's surface temperature T is about 3500 K. Its distance is about 200 pc, and from that and its brightness, its luminosity is estimated to be 140,000 times that of the Sun. What can we say about the size of Betelgeuse? Using the preceding equation, we can determine the following:

$$\frac{R_{\text{Betelgeuse}}}{R_{\text{Sun}}} = \sqrt{\frac{L_{\text{Betelgeuse}}}{L_{\text{Sun}}}} \times \frac{T^2_{\text{Sun}}}{T^2_{\text{Betelgeuse}}}$$

$$\frac{R_{\text{Betelgeuse}}}{R_{\text{Sun}}} = \sqrt{\frac{140{,}000}{1}} \times \frac{5{,}800^2}{3{,}500^2} = 374 \times 2.7 = 1{,}010$$

Betelgeuse has a radius more than 1,000 times larger than that of the Sun; such stars are called *supergiants*.

smallest types of stars, called white dwarfs, have radii that are only about 1 percent of the Sun's radius—about the size of Earth. The largest stars, called red supergiants, can have radii more than 1,000 times that of the Sun. There are many more stars toward the small end of this range—smaller than the Sun—than there are giant stars.

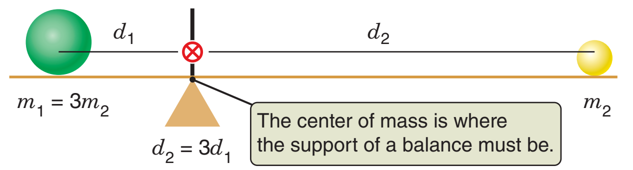


Figure 13.9 The center of mass of two objects is the "balance" point on a line joining the centers of two masses.

CHECK YOUR UNDERSTANDING 13.2

If star A has twice the surface temperature of the Sun but has the same luminosity as the Sun, the diameter of star A must be ________ the diameter of the Sun. (a) 16 times; (b) 4 times; (c) $\frac{1}{2}$; (d) $\frac{1}{4}$

13.3 Measuring the Masses of Stars in Binary Systems

Determining the mass of a star is difficult. Astronomers cannot use the amount of light from a star or the star's size as a measure of its mass. Stars can be large or small, faint or luminous. However, more massive stars *always* have stronger gravity. When astronomers are trying to determine the masses of astronomical objects, they almost always wind up looking for the effects of gravity. In Chapter 4, you learned that Kepler's laws of planetary motion are the result of gravity, and that the properties of the orbit of a planet can be used to measure the mass of the Sun. Similarly, astronomers can study two stars that orbit each other to determine their masses.

About half of the higher-mass stars in the sky are actually systems consisting of several stars moving about under the influence of their mutual gravity. Most of these are **binary stars** in which two stars orbit each other in elliptical orbits as predicted by Newton's version of Kepler's laws. This version of Kepler's laws can be used to find the mass of a star, as we will demonstrate. However, most low-mass stars are single, and low-mass stars far outnumber higher-mass stars, so most stars are single and their mass cannot be found this way. In this section, we will look at how astronomers measure the masses of stars in binary systems.

Binary Star Orbits

The **center of mass** is the balance point of a system. If the two objects were sitting on a seesaw in a gravitational field, the support of the seesaw would have to be directly under the center of mass for the objects to balance, as shown in **Figure 13.9**. When Newton applied his laws of motion to the problem of orbits, he found that two objects must move in elliptical orbits around each other, and that their common center of mass lies at one focus shared by both of the ellipses, as shown in **Figure 13.10**. The center of mass, which lies along the line between the two objects, remains stationary. The two objects will always be found on exactly opposite sides of the center of mass.

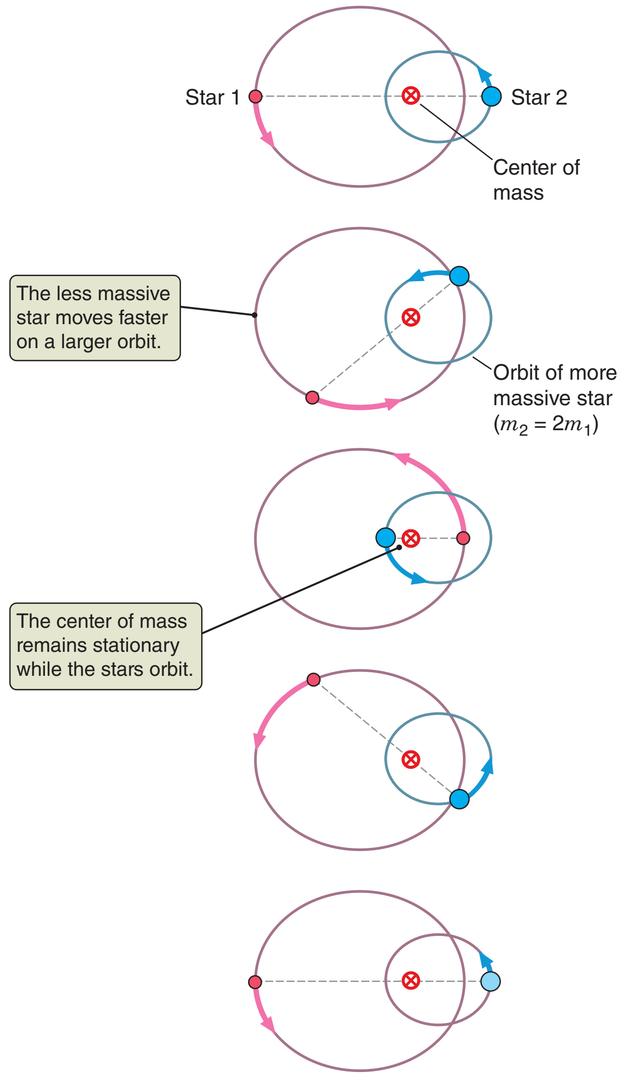


Figure 13.10 In a binary star system, the two stars orbit on elliptical paths about their common center of mass. In this case, the blue star has twice the mass of the red one. The eccentricity of the orbits is 0.5. There are equal time steps between the frames.

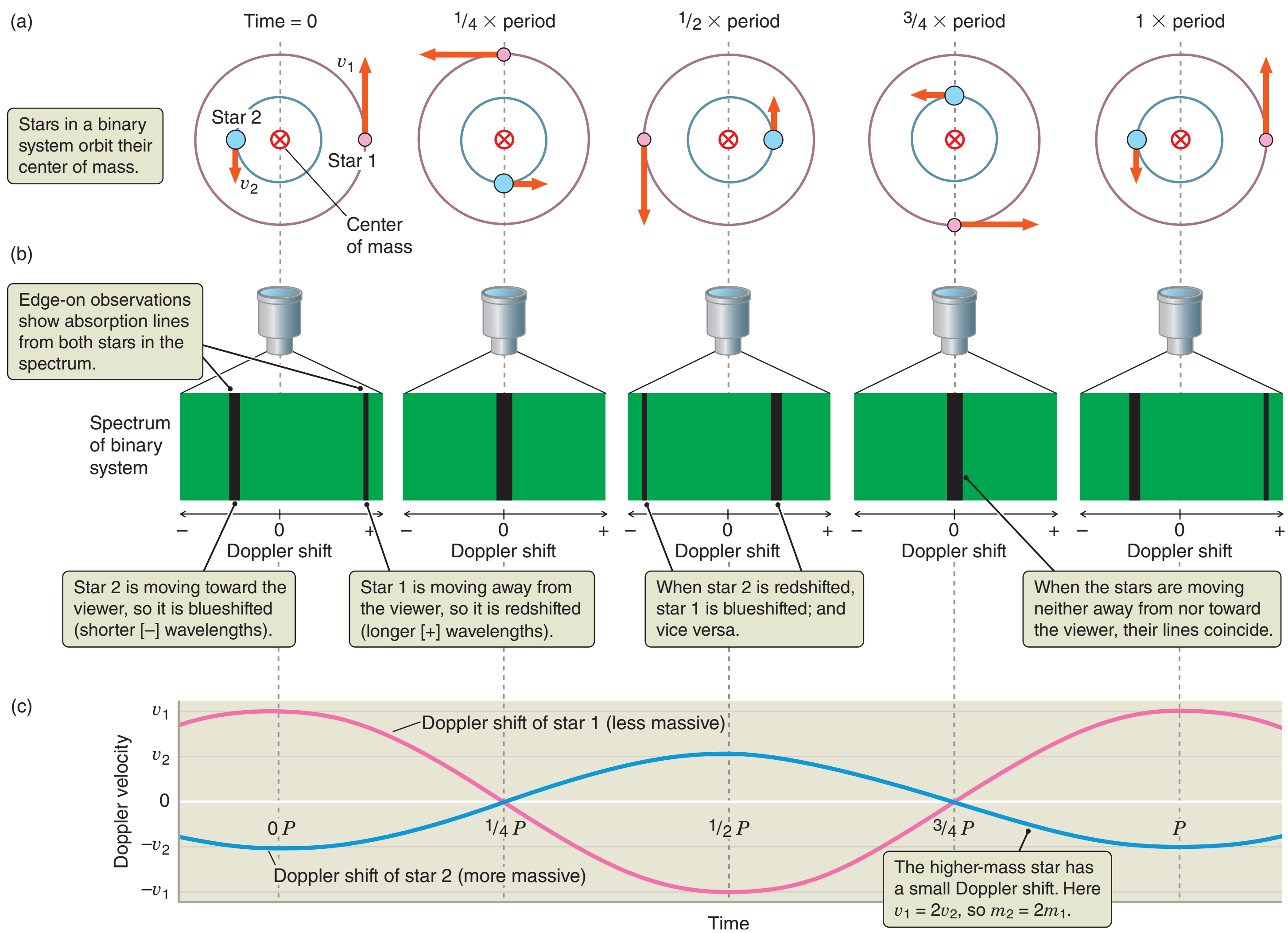


Figure 13.11 (a) The view from "above" the binary system shows that both stars orbit a common center of mass. (b) The spectrum of the combined system (seen edge-on) shows the spectral lines of each star shift back and forth. (c) Graphing the Doppler shift of star 1 with star 2 versus time reveals that star 2 has half the maximum Doppler shift, so star 2 is twice as massive as star 1. *P* is the period of the orbit.

Nebraska Simulation: Center of Mass Simulator

Imagine that you are watching a binary star as shown in **Figure 13.11a**. As seen from above, two stars orbit the common center of mass. Star 1, which is less massive, must complete its orbit in the same time as star 2, which is more massive. Because the less massive star has farther to go around the center of mass, it must be moving *faster* than the more massive star. In this view, no determination of the Doppler shift (Chapter 5) can be made because all the motion is in the plane of the sky, and none is toward or away from the observer.

When a system is edge-on to the observer, however, the observer can take advantage of the Doppler shift to find out about the motion. Figure 13.11b shows observations of the spectrum of the combined system associated with each position in Figure 13.11a. The spectral lines of the stars shift back and forth as they move toward and away from the observer. Because the two stars are always exactly on opposite sides of their center of mass, they are always moving in opposite directions. When star 2 approaches, star 1 recedes. The light coming from star 2 will be shifted to shorter wavelengths by the Doppler effect as it approaches, so the light will be blueshifted, and the light coming from star 1 will be shifted to longer wavelengths as it recedes, so the light will be redshifted. Half an orbital period later, the situation is reversed: lines from star 1 are blueshifted, and lines from star 2 are redshifted.

The less massive star has a larger orbit—and consequently moves more quickly—than the more massive star. Figure 13.11c compares the velocity obtained from the maximum Doppler shift for star 1 with the velocity obtained from maximum Doppler shift for star 2. This comparison gives the ratio of the masses of the two stars:

$$\frac{v_1}{v_2} = \frac{m_2}{m_1}$$

By observing the spectrum of the system, we can find from this equation the relative masses of the two stars; star 2 is two times as massive as star 1. But we can't find the actual mass of either star from these observations alone.

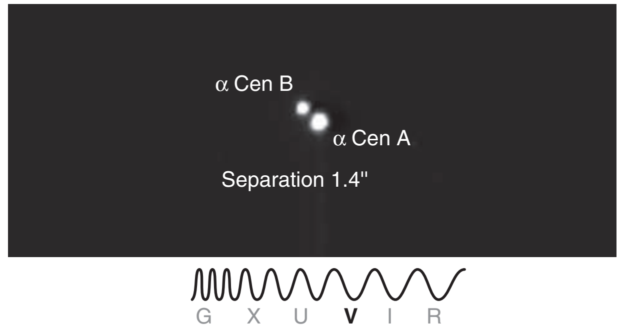


Figure 13.12 The two stars of a visual binary are resolved. These stars are two components of Alpha Centauri, the nearest star system to the Solar System.

The Masses of Binary Stars

Nebraska Simulation: Eclipsing Binary Simulator

In Chapter 4, we ignored the complexity of the motion of two objects around their common center of mass. Now, however, this very complexity enables us to measure the masses of the two stars in a binary system. If we can measure the period of the binary system and the average separation between the two stars, then Kepler's third law gives us the total mass in the system: the sum of the two masses. Because the analysis in the previous subsection gives us the ratio of the two masses, we now have two different relationships between two different unknowns. We have all we need to determine the mass of each star separately. In other words, if we know that star 2 is 2 times as massive as star 1, and we know that star 1 and star 2 together are 3 times as massive as the Sun, then we can calculate separate values for the masses of star 1 and star 2.

Depending on the type of system, there are several ways to measure the average separation and the orbital period. In a **visual binary** system, the system is close enough to Earth, and the stars are far enough from each other, that we can take pictures that show the two stars separately (**Figure 13.12**). Then, astronomers can directly measure the shapes and periods of the orbits of the two stars just by watching them as they orbit each other. These can be used with Doppler measurements of the radial (line-of-sight) velocities of the stars to solve for the ratio of the two masses.

In most binary systems, however, the two stars are so close together and far away from us that we cannot actually see the stars separately. The identification of these stars as binary systems is more indirect and comes from observing periodic variations in the *light* from the star or from observing periodic changes in the *spectrum* of the star. If we view a binary system nearly edge-on, so that one star passes in front of the other, it is called an **eclipsing binary**. An observer will see a repeating dip in brightness as one star passes in front of (eclipses) the other. If the stars are of different temperatures, there will be a repeating pattern of a smaller dip in brightness when the hotter star eclipses the cooler one, followed by a larger dip in brightness when the cooler star eclipses the hotter one, as shown in **Figure 13.13**. The pattern of these dips also gives an estimate of the relative sizes (radii) of the two stars. This procedure for identifying binary systems is similar to the transit method for finding extrasolar planets discussed in Chapter 7, and it works only when the system is viewed nearly edge-on. The Kepler space telescope

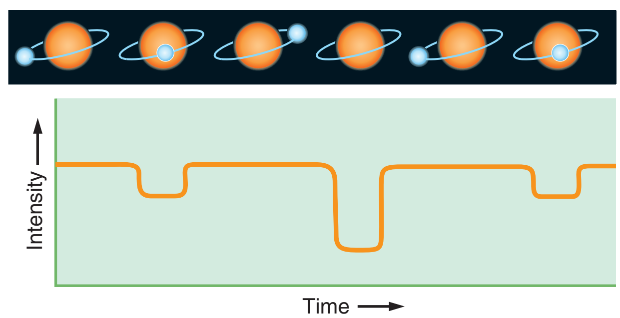


Figure 13.13 In an eclipsing binary system, the system is viewed nearly edge-on, so that the stars repeatedly pass behind one another, blocking some of the light. When the blue star passes in front of the larger, cooler star, less light is blocked than when it passes behind the red star. The shape of the dips in the light curve of an eclipsing binary can reveal information about the relative size and surface brightness of the two stars.

13.4 Working It Out Measuring the Mass of an Eclipsing Binary Pair

In Working It Out 4.3, we used Newton's version of Kepler's third law to calculate the mass of the Sun by observing the orbital period of one of its planets. In that special case, the Sun's mass is so much greater than the mass of a planet that the planet's mass is negligible. In the case of two stars, however, the masses are similar, so neither is negligible, and we need to keep both in the equations. Newton showed that if two objects with masses m_1 and m_2 are in orbit about each other, then the period of the orbit, P, is related to the average distance between the two masses, the semimajor axis A, by the equation

$$P^2 = \frac{4\pi^2 A^3}{G(m_1 + m_2)}$$

Rearranging this equation a bit turns it into an expression for the sum of the masses of the two objects:

$$m_1 + m_2 = \frac{4\pi^2}{G} \times \frac{A^3}{P^2}$$

We could use the equation this way, with the masses of the two stars in kilograms (kg), the distance between them in kilometers, the period of their orbit in seconds, and the gravitational constant G in kilometers, kilograms, and seconds. However, astronomers often think about stellar masses in units of the Sun's mass. If we divide this equation by the "mass of the Sun" equation in Working It Out 4.3,

$$M_{\text{Sun}} = \frac{4\pi^2}{G} \times \frac{A^3}{P^2}$$

(where M_{Sun} = mass of the Sun, $A = 1$ AU, and $P = 1$ year), then the constants cancel out and this equation simplifies to

$$\frac{m_1}{M_{\text{Sun}}} + \frac{m_2}{M_{\text{Sun}}} = \frac{A^3_{\text{AU}}}{P^2_{\text{years}}}$$

Therefore, if we know both m_1/m_2 from measuring velocities by Doppler shifts and $m_1 + m_2$ from the observed orbital properties, we can solve for the separate values of m_1 and m_2.

Suppose you are an astronomer studying a binary star system. After observing the star for several years, you accumulate the following information about the system:

1. The star is an eclipsing binary.
2. The period of the orbit is 2.63 years.
3. Star 1 has a Doppler velocity that varies between +20.4 and −20.4 km/s.
4. Star 2 has a Doppler velocity that varies between +6.8 and −6.8 km/s.
5. The stars are in circular orbits. You know this because the Doppler velocities about the star are symmetric; the approach and recession speeds of the star are equal.

These data are summarized in **Figure 13.14.** You begin your analysis by noting that the star is an eclipsing binary, which tells you that the orbit of the star is edge-on to your line of sight. The Doppler velocities tell you the total orbital velocity of each star, and you determine the size of the orbits using the relationship

$$\text{Distance} = \text{Speed} \times \text{Time}$$

In one orbital period, star 1 travels around a circle—a distance of

$$d = (20.4\ \text{km/s}) \times (2.63\ \text{yr}) = 53.7\ \text{km} \times \text{yr/s}$$

Multiply by the number of seconds in a year:

$$d = 53.7 \frac{\text{km} \times \text{yr}}{\text{s}} \times \frac{3.16 \times 10^7\ \text{s}}{\text{yr}} = 1.70 \times 10^9\ \text{km}$$

This distance is the circumference of the star's orbit, or 2π times the radius of the star's orbit, A_1. Thus, star 1 is following an orbit with a radius of

$$A_1 = \frac{d}{2\pi} = \frac{1.70 \times 10^9\ \text{km}}{2\pi} = 2.7 \times 10^8\ \text{km}$$

has been observing and discovering thousands of eclipsing binaries in addition to finding new extrasolar planets.

If a binary system is a **spectroscopic binary**, the spectral lines of the two stars show periodic changes as they are Doppler-shifted away from each other, first in one direction and then in the other, as shown in Figure 13.11. The period of the orbit is determined from the time it takes for a set of spectral lines to go from approaching to receding and back again. The orbital velocities of the stars and the period of the orbit give the size of the orbit because distance equals velocity multiplied by time. Consequently, astronomers can estimate the combined masses of the two stars. To calculate the individual masses, an estimate of the tilt of the orbit is needed. Thus, spectroscopic binary masses are more approximate than those in eclipsing binary systems.

To convert this to astronomical units, use the relation 1 AU = 1.50×10^8 km:

$$A_1 = 2.7 \times 10^8 \text{ km} \times \frac{1 \text{ AU}}{1.50 \times 10^8 \text{ km}} = 1.8 \text{ AU}$$

A similar analysis of star 2 shows that its orbit has a radius of $A_2 =$ 0.6 AU.

Next, apply Newton's version of Kepler's third law. Because the stars are always on opposite sides of the center of mass, $A_{AU} = 1.8$ AU + 0.6 AU = 2.4 AU. Because you know A and the period P (measured as 2.63 years), you can calculate the total mass of the two stars:

$$\frac{m_1}{M_{Sun}} + \frac{m_2}{M_{Sun}} = \frac{(A_{AU})^3}{(P_{years})^2} = \frac{(2.4)^3}{(2.63)^2} = 2.0$$

So you have learned that the combined mass of the two stars is twice the mass of the Sun. To sort out the individual masses of the stars, use the measured velocities and the fact that the mass and velocity are inversely proportional:

$$\frac{m_2}{m_1} = \frac{v_1}{v_2} = \frac{20.4 \text{ km/s}}{6.8 \text{ km/s}} = 3.0$$

Star 2 is 3 times as massive as star 1. In mathematical terms, $m_2 = 3 \times m_1$. Substituting into the equation

$$m_1 + m_2 = 2.0\,M_{Sun}$$

gives

$$m_1 + 3m_1 = 2.0\,M_{Sun}$$

or $4m_1 = 2.0\,M_{Sun}$, so $m_1 = 0.5\,M_{Sun}$. Because $m_2 = 3 \times m_1$, then $m_2 =$ $1.5\,M_{Sun}$.

Star 1 has a mass of $0.5\,M_{Sun}$, and star 2 has a mass of $1.5\,M_{Sun}$. You have just found the masses of two distant stars.

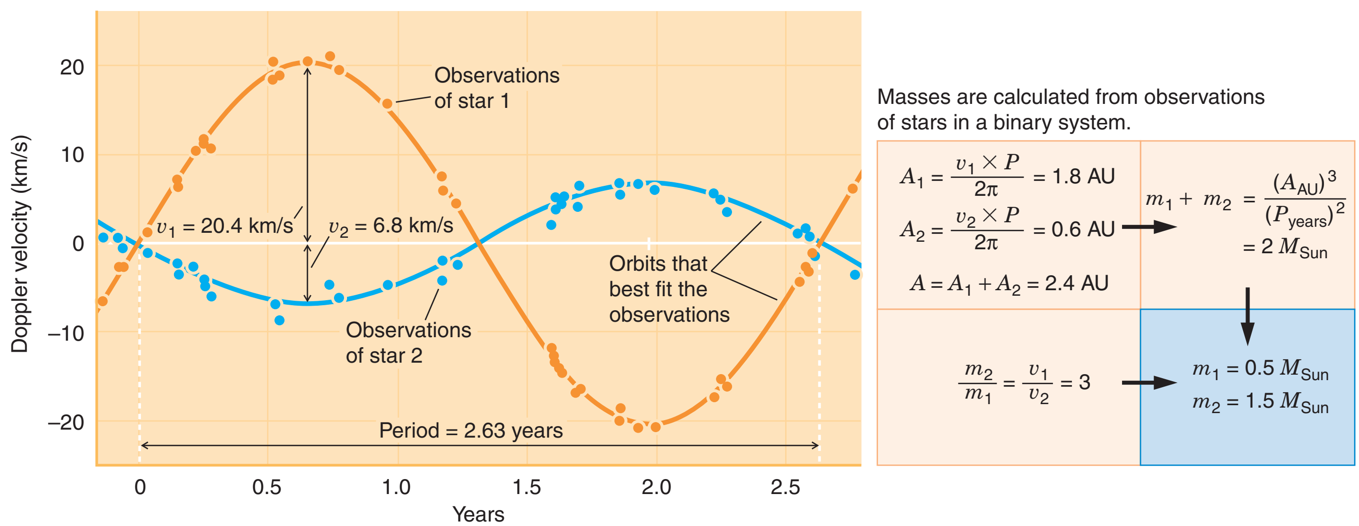


Figure 13.14 Doppler velocities of the stars in an eclipsing binary are used to measure the masses of the stars.

A binary system can fall into more than one of these three categories, regardless of how it was originally discovered. If a spectroscopic binary system is also a visual or eclipsing binary, then the orbit and masses of the stars can be completely solved (**Working It Out 13.4**). Historically, most stellar masses were measured for stars in eclipsing binary systems, rather than for those in visual or spectroscopic binaries. But new observational capabilities have increased the number of known visual binaries by greatly improving the ability to see the stars in a binary directly. Accurate measurements of masses have been obtained for several hundred binary stars, about half of which are eclipsing binaries. The range of stellar masses found in this way is not nearly as great as the range of stellar luminosities. The least massive stars have masses of about 0.08 M_{Sun}; the most massive stars appear to have masses up to about 200 M_{Sun}.

CHECK YOUR UNDERSTANDING 13.3

Which of the following properties must be measured to determine the masses of stars in a typical binary system? (Choose all that apply.) (a) the period of the orbits of the two stars; (b) the average separation between the two stars; (c) the radii of the two stars; (d) the velocities of the two stars.

13.4 The Hertzsprung-Russell Diagram Is the Key to Understanding Stars

Nebraska Simulation: Hertzsprung-Russell Diagram Explorer

Measuring some basic properties of stars is only the first step in understanding; the next step is to look for patterns in their properties. In the early part of the 20th century, Ejnar Hertzsprung (1873–1967) and Henry Norris Russell (1877–1957) independently studied the properties of stars. Both astronomers plotted the

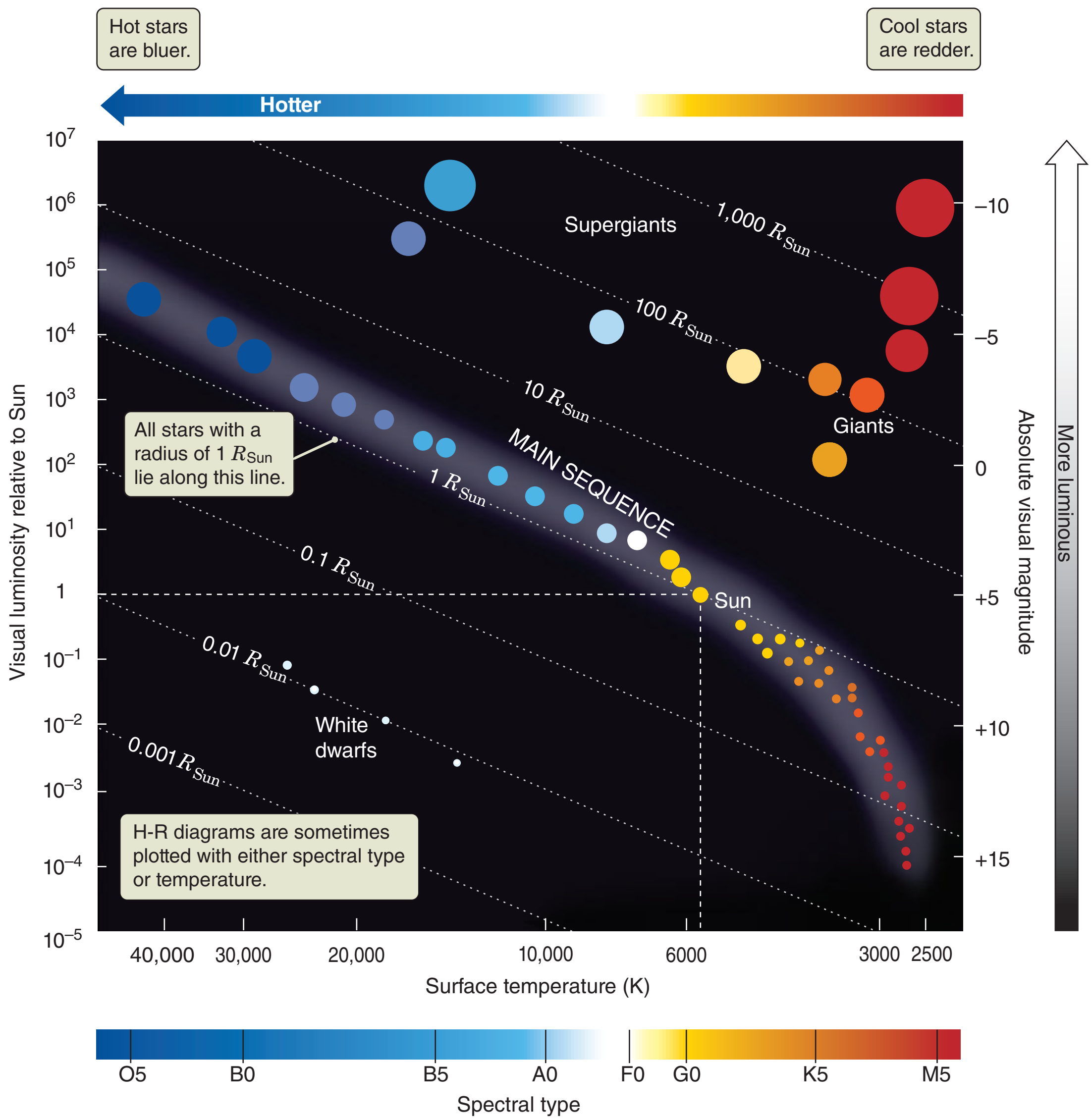


Figure 13.15 The Hertzsprung-Russell, or H-R, diagram is used to plot the properties of stars. More luminous stars are at the top of the diagram. Hotter stars are on the left. Stars of the same radius (*R*) lie along the dotted lines moving from upper left to lower right. Absolute magnitudes are discussed in Working It Out 13.2 and Appendix 7.

luminosities of stars versus their surface temperatures—a diagram that came to be known as the Hertzsprung-Russell diagram, or simply **H-R diagram**. We will use this diagram often for the study of stars. In this section, we take a first look at this important diagram and the way stars are organized within it.

The H-R Diagram

AstroTour: H-R Diagram

We begin with the layout of the H-R diagram, shown in **Figure 13.15**. The spectral type is plotted on the horizontal axis (the *x*-axis), along with the surface temperature plotted backward: temperature is higher on the left and lower on the right. Hot blue stars are on the left side of the H-R diagram; cool red stars are on the right. Temperature is plotted logarithmically. This means that the size of an interval along the axis from a point representing a star with a surface temperature of 40,000 K to one with a surface temperature of 20,000 K—a temperature change by a factor of 2—is the same as the size of an interval between points representing a star with a temperature of 10,000 K and a star with a temperature of 5000 K, which is also a temperature change by a factor of 2. The horizontal axis is sometimes labeled with another characteristic that corresponds to temperature, such as the color.

The luminosity of stars is plotted along the vertical axis (the *y*-axis). More luminous stars are toward the top of the diagram, and less luminous stars are toward the bottom. Sometimes, the luminosity axis is labeled with the absolute visual magnitude instead of luminosity, as shown on the right-hand *y*-axis. As with the temperature axis, luminosities are plotted logarithmically. In this case, each step along the left-hand *y*-axis corresponds to a multiplicative factor of 10 in the luminosity. To understand why the plotting is done this way, recall that the most luminous stars are 10 billion times more luminous than the least luminous stars, yet all of these stars must fit on the same plot.

Each point on the H-R diagram is specified by a surface temperature and luminosity. Therefore, we can use the luminosity-temperature-radius relationship described earlier in the chapter to find the radius of a star at that point as well. A star in the upper right corner of the H-R diagram is very cool, so each square meter of its surface radiates only a small amount of energy. But this star is also extremely luminous. It must be huge to account for its high luminosity, despite the feeble radiation coming from each square meter of its surface. Conversely, a star in the lower left corner of the H-R diagram is very hot, which means that a large amount of energy is coming from each square meter of its surface. However, this star has a very low overall luminosity, so it must be very small. Moving up and to the right takes you to larger and larger stars. Moving down and to the left takes you to smaller and smaller stars. All stars of the same radius lie along slanted lines across the H-R diagram. Astronomers can note the properties of a star—its temperature, color, size, and luminosity—from a glance at its position on the H-R diagram. The discovery and study of these patterns led to an understanding of the astrophysics of stars (**Process of Science Figure**).

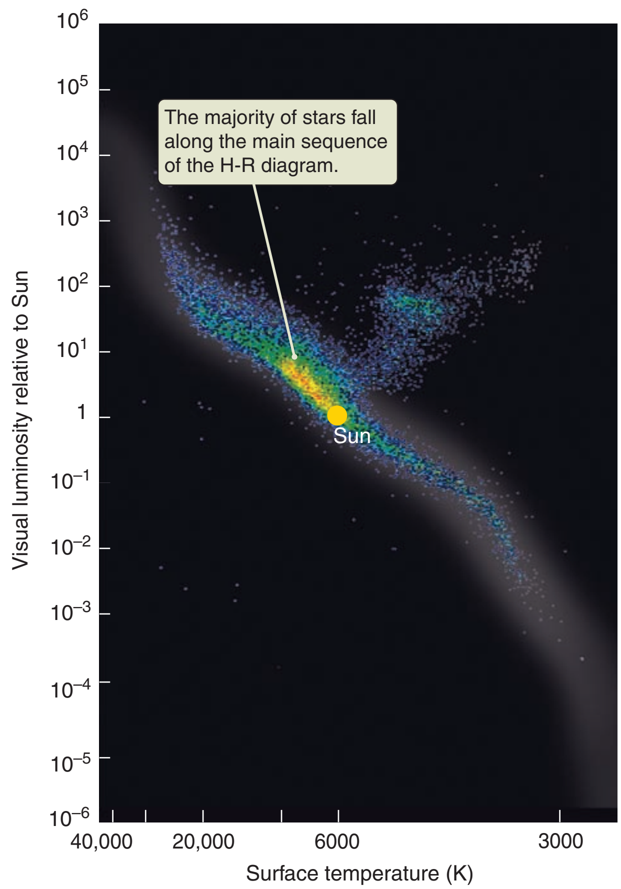


Figure 13.16 An H-R diagram for 16,600 stars plotted from data obtained by the Hipparcos satellite clearly shows the main sequence. Most of the stars lie along this band running from the lower right of the diagram toward the upper left.

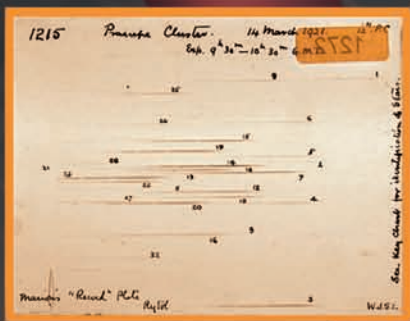
The Main Sequence

Figure 13.16 shows 16,600 nearby stars plotted on an H-R diagram. The data are based on observations of stars near enough for parallax measurements obtained by the Hipparcos satellite. A quick look at this diagram immediately

Process of Science

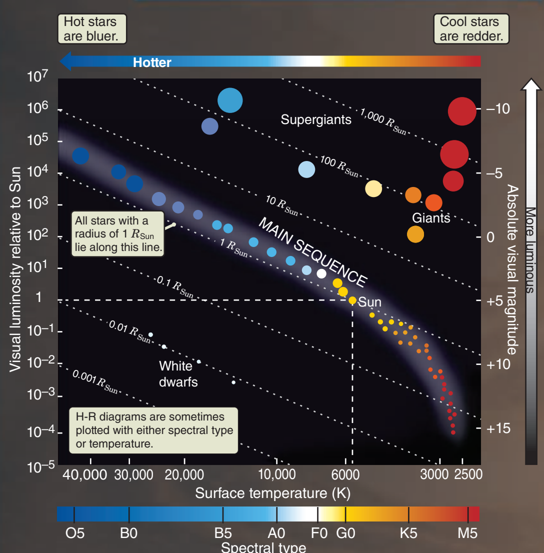
SCIENCE IS COLLABORATIVE

An understanding of the meaning behind stellar data took decades, and the contributions of dozens of people, all working toward a common goal.



The Observations (~1880s):
About 500,000 photographs of stellar spectra are obtained, by many astronomers at many telescopes.

The Classification (~1900s):
Annie Jump Cannon leads a team that classifies all the spectra according to the strengths of particular absorption lines at particular wavelengths.



The Graph (~1910s):
Hertzsprung and Russell independently develop what will later be called the H-R diagram. They do not understand why the *x*-axis, when ordered O-B-A-F-G-K-M, gives such a nice band across the middle. Russell hypothesizes this must come from a single stellar characteristic.

The Understanding (~1920s):
Meghnad Saha shows that the stellar characteristic in question is temperature. Cecilia Payne-Gaposchkin shows that stars are mostly composed of hydrogen and helium. Modern astrophysics is born; others go on to develop the understanding of stellar atmospheres.

Scientific discoveries sometimes seem to occur suddenly. However, new scientific knowledge is usually the effort of many minds working for years to solve a problem.

shows what was first discovered in the original diagrams of Hertzsprung and Russell. About 90 percent of the stars in the sky lie in a well-defined region running across the H-R diagram from lower right to upper left, known as the **main sequence**. On the left end of the main sequence are the O stars: hotter, larger, and more luminous than the Sun. On the right end of the main sequence are the M stars: cooler, smaller, and fainter than the Sun. If you know where a star lies on the main sequence, then you know its approximate luminosity, surface temperature, and size.

The H-R diagram supplies a useful method for finding the distance to main-sequence stars. Astronomers can determine whether a star is on the main sequence by looking at the absorption lines in its spectrum. The spectral type is also determined from the spectral lines, and this spectral type indicates the star's temperature. Once this value on the *x*-axis is known, we can then read up to the main sequence and then across to the *y*-axis to find the star's luminosity. Recall that the luminosity, brightness, and distance are all connected. We can think about how far away a star of a particular luminosity must be to have the observed brightness. So we can find the star's distance by comparing a star's luminosity, obtained from the H-R diagram, with its apparent brightness. This method of determining distances to main-sequence stars from the spectra, luminosity, and brightness of stars is called **spectroscopic parallax**. Details of this method are discussed in Appendix 7. Despite the similarity between the names, this method is very different from the parallax method using trigonometry. Spectroscopic parallax is useful to much larger distances than trigonometric parallax, although it is less precise.

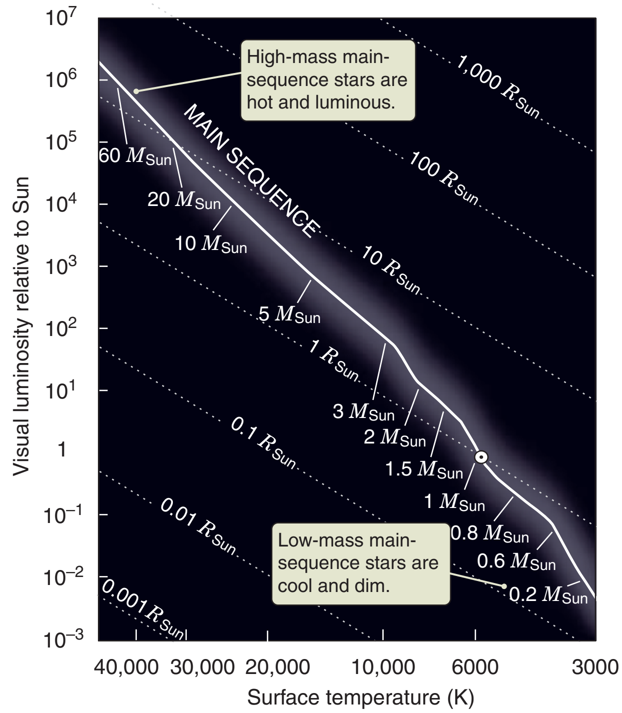


Figure 13.17 Mass determines the location of a star along the main sequence..

Nebraska Simulation: Spectroscopic Parallax Simulator

From a combination of observations of binary star masses, parallax, luminosity measurements, and mathematical models, astronomers have determined that stars of different masses lie on different parts of the main sequence. Stellar mass increases smoothly from the lower right to the upper left along the main sequence. If a main-sequence star is less massive than the Sun, it is also smaller, cooler, redder, and less luminous than the Sun, and it is located to the lower right of the Sun on the main sequence. Conversely, if a main-sequence star is more massive than the Sun, it is also larger, hotter, bluer, and more luminous than the Sun, and it is located to the upper left of the Sun on the main sequence, as illustrated in **Figure 13.17**. The mass of a star determines where on the main sequence the star will lie.

Table 13.2 summarizes the properties of the different spectral classes of main-sequence stars. *All* main-sequence stars with a mass of 1 M_{Sun} are G2 stars like the Sun and have the *same* surface temperature, size, and luminosity as the Sun. Similarly, if a main-sequence star is classified as B0, it has these properties: a surface temperature of about 30,000 K, a luminosity about 32,500 times that of the Sun, a mass of about 17.5 M_{Sun}, and a radius of about 6.7 R_{Sun}. If a different main-sequence star is classified as M5, it has a surface temperature of 3,170 K, a luminosity of about 0.008 L_{Sun}, a mass of about 0.21 M_{Sun}, and a radius of about 0.29 R_{Sun}.

The relationship between the mass and the luminosity of stars is very sensitive. Relatively small differences in the masses of stars result in large differences in their main-sequence luminosities. From determining the luminosities of binary stars with measured mass, a relationship between the mass and luminosity was found. This

TABLE 13.2 The Properties of Main Sequence Stars

Spectral Type	Temperature (K)	Mass (M_{Sun})	Radius (R_{Sun})	Luminosity (L_{Sun})
O5	42,000	60	13	500,000
B0	30,000	17.5	6.7	32,500
B5	15,200	5.9	3.2	480
A0	9800	2.9	2.0	39
A5	8200	2.0	1.8	12.3
F0	7300	1.6	1.4	5.2
F5	6650	1.4	1.2	2.6
G0	5940	1.05	1.06	1.25
G2 (Sun)	5780	1.00	1.00	1.0
G5	5560	0.92	0.93	0.8
K0	5150	0.79	0.93	0.55
K5	4410	0.67	0.80	0.32
M0	3840	0.51	0.63	0.08
M5	3170	0.21	0.29	0.008

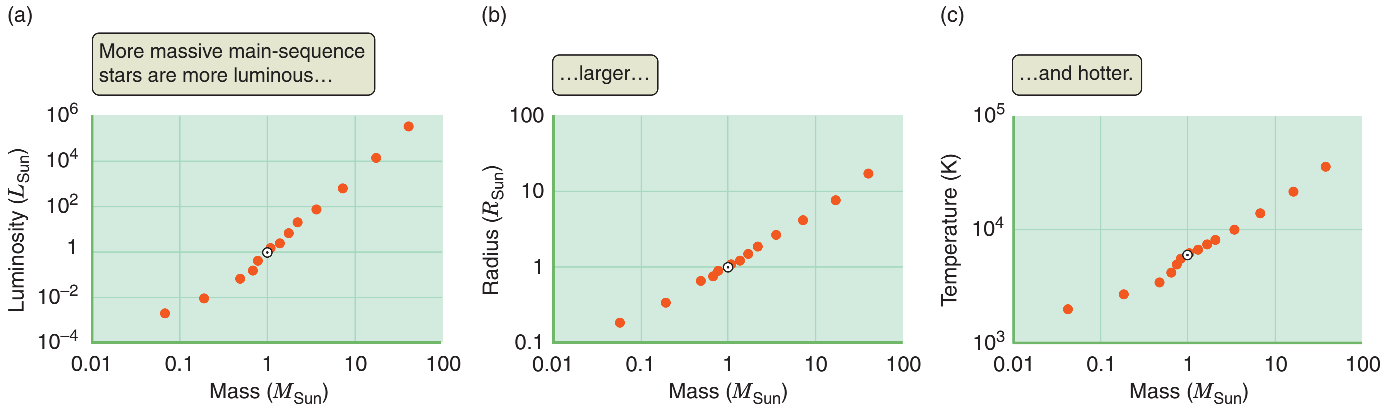


Figure 13.18 These graphs plot luminosity (a), radius (b), and temperature (c) versus mass for stars along the main sequence. The mass of a main-sequence star determines its other properties.

mass-luminosity relationship, usually expressed as $L \propto M^{3.5}$, is shown in **Figure 13.18a**. The exact exponent varies somewhat for different ranges of stellar masses, but this method is useful for estimating masses of single stars. The mass also correlates to the size of a star, shown in Figure 13.18b, and to the temperature of a star, shown in Figure 13.18c.

The mass and chemical composition of a main-sequence star determine its other characteristics: how large it is, what its surface temperature is, how luminous it is, what its internal structure is, how long it will live, how it will evolve, and what its final fate will be. A star must have a balance between gravity trying to hold the star together and the energy released by nuclear reactions in the interior of the star trying to blow it apart. The mass of the star determines the strength of its gravity, which in turn determines how much energy must be generated in its interior to prevent it from collapsing under its own weight. The mass of a star determines where the balance is struck.

Stars Not on the Main Sequence

Although 90 percent of stars are main-sequence stars, some stars are found in the upper right portion of the H-R diagram, well above the main sequence (see Figure 13.15). So they must be luminous, cool, and large, with radii hundreds or thousands of times the radius of the Sun. These stars are called giants. At the other extreme are stars found in the far lower left corner of the H-R diagram. These stars are the tiny white dwarfs, comparable to the size of Earth. Their small surface areas explain why they have such low luminosities, despite having high temperatures.

Stars that lie off the main sequence on the H-R diagram can be identified by their luminosities (determined by their distance) or by slight differences in their spectral lines. The width of a star's spectral lines is an indicator of the density and surface pressure of gas in the star's atmosphere. In general, denser stars have broader lines. Puffed-up stars above the main sequence have lower densities and lower surface pressure and narrower absorption lines compared to main-sequence stars.

When using the H-R diagram to estimate the distance to a star by the spectroscopic parallax method, astronomers must know whether the star is on, above, or below the main sequence in order to find the star's luminosity. The spectral line widths of stars both on and off the main sequence indicate **luminosity class**, which tells us the relative *size* of the star within each spectral class. Supergiant stars, which are the largest stars that we see, are luminosity class I, bright giants are class II, giants are class III, subgiants are class IV, main-sequence stars are class V, and white dwarfs are class WD. Luminosity classes I–IV lie above the main sequence, while class WD falls below and to the left of the main sequence, as shown in **Figure 13.19**. Thus, the complete spectral classification of a star includes both its spectral type (which indicates temperature and color) and its luminosity class (which indicates relative size).

The existence of the main sequence, together with the fact that the mass of a main-sequence star determines where on the sequence it will lie, is a grand pattern that points to the possibility of a deep understanding of stars. The existence of stars that do *not* follow this pattern raises yet more questions. In the coming chapters, you will learn that the main sequence tells us what stars are and how they work, and that stars off the main sequence reveal how stars form, how they evolve, and how they die. **Table 13.3** summarizes the techniques that astronomers use to determine some of the basic properties of stars. Of the properties listed in the table, only temperature, distance, and composition can be *measured*. Luminosity must be *inferred* from the H-R diagram or calculated from distance and brightness, and size and mass must be *calculated*. Other properties that can be measured include brightness, color, spectral type, and parallax shift.

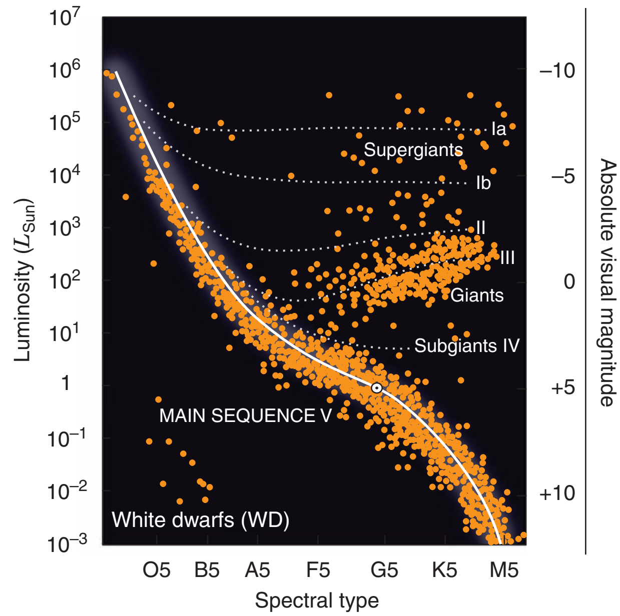


Figure 13.19 Stellar luminosity classes indicate the size (radius) of a star at each spectral type.

CHECK YOUR UNDERSTANDING 13.4

Choose the two qualities that describe a star located in the lower right of the H-R diagram: (a) hot; (b) cold; (c) high luminosity; (d) low luminosity.

TABLE 13.3 Taking the Measure of Stars

Property	Methods
Luminosity	• For a star with a known distance, measure the brightness, then apply the inverse square law of radiation: $$\text{Luminosity} = 4\pi \times \text{Distance}^2 \times \text{Brightness}$$ • For a star without a known distance, take a spectrum of the star to determine its spectral and luminosity classes, plot them on an H-R diagram, and read the luminosity from the diagram.
Temperature	• Measure the color index of the star using blue and visual filters. Use Wien's law to relate the color to a temperature. • Take a spectrum of the star, and estimate the temperature from its spectral class by noting which spectral lines are present.
Distance	• For a relatively nearby star (within a few hundred parsecs), measure the parallax shift of the star over the course of the year. • For a more distant star, find the luminosity using the H-R diagram as noted earlier, and then use the spectroscopic parallax method to relate luminosity, distance, and brightness.
Size	• For a few of the largest and closest stars, measure the size directly or by the length of eclipse in eclipsing binary stars. • From the width of the star's spectral lines, estimate the luminosity class (supergiant, giant, or main sequence). • For a star with known luminosity and temperature, use the Stefan-Boltzmann law to calculate the star's radius (the luminosity-temperature-radius relationship).
Mass	• Measure the motions of the stars in a binary system, and use these to determine the orbits of the stars, then apply Newton's form of Kepler's third law. • For a non-binary star, use the mass-luminosity relationship to estimate the mass from the luminosity.
Composition	• Analyze the lines in the star's spectrum to measure chemical composition.

Origins

Habitable Zones

How might a basic property of a star, such as its luminosity, color, mass, or surface temperature, affect the chance of there being a planet with life in orbit around that star? The only known life is that on planet Earth, where liquid water was essential for its formation and evolution. Whether liquid water is an absolute requirement for life elsewhere is not known, but the presence of water is a good starting point for determining where to look. So astronomers look for planets that are at the right distance from their stars to have a planetary temperature that permits water to exist in a liquid state on their surfaces—a range of distances known as the **habitable zone**. On planets that lie inside the habitable zone, water would exist only as a vapor—if at all. On planets that lie outside the habitable zone, water would be permanently frozen as ice.

Recall from the Chapter 5 "Origins" and Working It Out 5.4 that an important factor for estimating the temperature of a planet is the brightness of the sunlight that falls on that planet. This factor depends on the luminosity of the star and the planet's distance from the star. In the Solar System, the habitable zone ranges from about ~0.9 to ~1.4 AU, which includes Earth but just misses Venus and Mars. Main-sequence stars that are less luminous than the Sun are cooler and have narrower habitable zones, minimizing the chance that a planet will form within that slender zone. Main-sequence stars that are more massive than the Sun are hotter and have larger habitable zones. **Figure 13.20** illustrates these zones around Sun-like, hotter, and cooler stars.

In the past few years, astronomers have started to find planets in the habitable zones of their respective stars. Methods of planet detection, as discussed in Chapter 7, work best when the planet is close to its star. Using the transit method, the Kepler Mission has identified and confirmed 1–2 dozen planets in habitable zones as of this writing and has found more candidate planets that need to be confirmed.

The distance from a star at a certain temperature is not the only consideration for whether a planet has water. The presence of a planetary atmosphere is also a factor. More massive planets can retain their atmospheres, which can trap heat and raise the planet's temperature, as we saw in Chapter 5 for Venus and Earth. Smaller planets have a lower gravitational pull and may not be able to keep an atmosphere. Additionally, some habitable zones may be near planets, not stars. Some of the giant planets in the cold outer part of our own Solar System have moons with liquid water. The heat keeping the water liquid comes from the nearby planet, not from the Sun.

Finally, we note that *habitable* does not mean *inhabited*; it only means the planet is at the right distance from its star that it could have liquid water. Identifying planets in their habitable zone is a first step to selecting which planets are most interesting for further study.

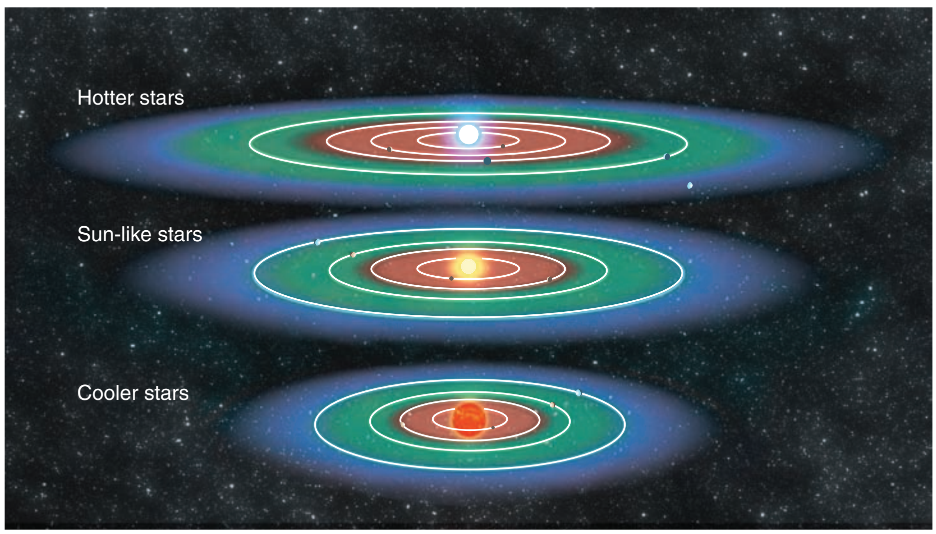


Figure 13.20 The distance and extent of a habitable zone (green) surrounding a star depends on the star's temperature. Regions too close to the star are too hot (red) and those too far away are too cold (blue) to permit the existence of liquid water. The orbits of Mercury, Venus, Earth, and Mars have been drawn around these stars for scale.

ARTICLES QUESTIONS

NASA reports on a new technique for using the Hubble Space Telescope to measure parallax angles.

NASA's Hubble Extends Stellar Tape Measure 10 Times Farther into Space

NASA Press Release

Using NASA's Hubble Space Telescope, astronomers now can precisely measure the distance of stars up to 10,000 light-years away—10 times farther than previously possible.

Astronomers have developed yet another novel way to use the 24-year-old space telescope by employing a technique called spatial scanning, which dramatically improves Hubble's accuracy for making angular measurements. The technique, when applied to the age-old method for gauging distances called astronomical parallax, extends Hubble's tape measure 10 times farther into space (**Figure 13.21**).

"This new capability is expected to yield new insight into the nature of dark energy, a mysterious component of space that is pushing the universe apart at an ever-faster rate," said Nobel laureate Adam Riess of the Space Telescope Science Institute (STScI) in Baltimore, Maryland.

Parallax, a trigonometric technique, is the most reliable method for making astronomical distance measurements, and a practice long employed by land surveyors here on Earth. The diameter of Earth's orbit is the base of a triangle, and the star is the apex where the triangle's sides meet. The lengths of the sides are calculated by accurately measuring the three angles of the resulting triangle.

Astronomical parallax works reliably well for stars within a few hundred light-years of Earth. For example, measurements of the distance to Alpha Centauri, the star system closest to our Sun, vary only by 1 arcsecond. This variance in distance is equal to the apparent width of a dime seen from 2 miles away.

Stars farther out have much smaller angles of apparent back-and-forth motion that are extremely difficult to measure. Astronomers have pushed to extend the parallax yardstick ever deeper into our galaxy by measuring smaller angles more accurately.

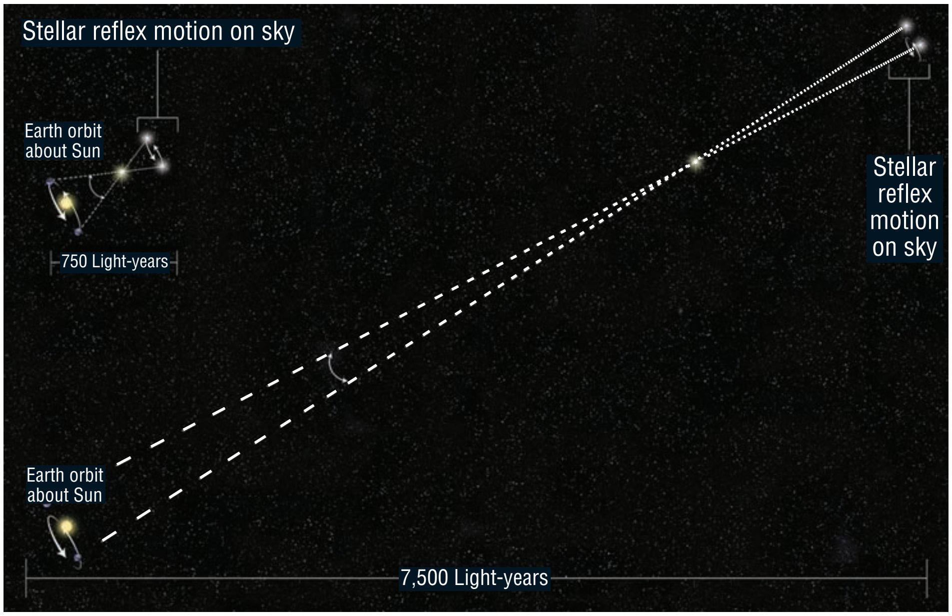


Figure 13.21 By applying a technique called spatial scanning to an age-old method for gauging distances called astronomical parallax, scientists now can use NASA's Hubble Space Telescope to make precision distance measurements 10 times farther into our galaxy than previously possible.

This new long-range precision was proven when scientists successfully used Hubble to measure the distance of a special class of bright stars called Cepheid variables, approximately 7,500 light-years away in the northern constellation Auriga. The technique worked so well, they are now using Hubble to measure the distances of other far-flung Cepheids.

Such measurements will be used to provide firmer footing for the so-called cosmic "distance ladder." This ladder's "bottom rung" is built on measurements to Cepheid variable stars that, because of their known brightness, have been used for more than a century to gauge the size of the observable universe. They are the first step in calibrating far more distant extra-galactic milepost markers such as Type Ia supernovae.

Riess and the Johns Hopkins University in Baltimore, Maryland, in collaboration with Stefano Casertano of STScI, developed a technique to use Hubble to make measurements as small as five-billionths of a degree.

To make a distance measurement, two exposures of the target Cepheid star were taken 6 months apart, when Earth was on opposite sides of the Sun. A very subtle shift in the star's position was measured to an accuracy of 1/1,000 the width of a single image pixel in Hubble's Wide Field Camera 3, which has 16.8 megapixels total. A third exposure was taken after another 6 months to allow for the team to subtract the effects of the subtle space motion of stars, with additional exposures used to remove other sources of error.

Riess shares the 2011 Nobel Prize in Physics with another team for his leadership in the 1998 discovery that the expansion rate of the universe is accelerating—a phenomenon widely attributed to a mysterious, unexplained dark energy filling the universe. This new high-precision distance measurement technique is enabling Riess to gauge just how much the universe is stretching. His goal is to refine estimates of the universe's expansion rate to the point where dark energy can be better characterized.

ARTICLES QUESTIONS

1. How many parsecs are in 7,500 light-years? in 10,000 light-years?
2. What is the parallax angle of a star that is 7,500 light-years away? 10,000 light-years away?
3. Why were the exposures taken 6 months apart?
4. Why is it better to make parallax measurements from space than from the ground on Earth?
5. How does improving the accuracy of the distances to nearby stars from trigonometric parallax affect astronomers' estimates of the distances to farther stars using spectroscopic parallax?

Summary

Finding the distances to stars is a difficult but important task for astronomers. Parallax and spectroscopic parallax are two of the methods that astronomers use to determine distances to stars. Brightness and distance can be used to obtain the luminosity. Careful study of the light from a star, including its spectral lines, gives the temperature, size, and composition of the star. Study of binary systems gives the masses of stars of various spectral types, which we can extend to all stars of the same spectral type. The H-R diagram shows the relationship among the various physical properties of stars. The mass of a star will be the major determining factor in its changes over time. The habitable zone is the distance from a star in which a planet could have the right temperature for liquid water to exist on its surface. Stars of different luminosities and temperatures have habitable zones of different widths at different distances from the star.

LG 1 **Explain how the brightness of nearby stars and their distances from Earth are used to determine how luminous they are.** The distance to a nearby star is measured by finding its parallax—by measuring how the star's apparent position changes in the sky over the course of a year. The nearest star (other than the Sun) is about 4 light-years (1.3 parsecs) away. The brightness of a star in the sky can be measured directly, and brightness and distance can be used to obtain the star's luminosity—how much light the star emits.

LG 2 **Explain how astronomers obtain the temperatures, sizes, and composition of stars.** The temperature of a star is determined by its color, with blue stars being hotter and red stars being cooler. The radius can be computed from the temperature and luminosity of the star. Small, cool stars greatly outnumber large, hot stars. Spectral lines carry a great deal of information about a star, including what chemical elements and molecules are present in the star.

LG 3 **Describe how astronomers estimate the masses of stars.** Masses of stars are measured in binary star systems by observing the effects of the gravitational pull between the stars. Newton's universal law of gravitation and Kepler's laws connect the motion of the star to the forces they experience, and thus to their masses.

LG 4 **Classify stars, and organize this information on a Hertzsprung-Russell (H-R) diagram.** The H-R diagram shows the relationship among the various physical properties of stars. Temperature increases to the left, so that hotter stars lie on the left side of the diagram, while cooler stars lie on the right. Luminosity increases vertically, so that the most luminous stars lie near the top of the diagram. A star's luminosity class and temperature indicate its size. The mass and composition of a main-sequence star determine its luminosity, temperature, and size. Ninety percent of stars lie along the main sequence.

LG 5 **Explain how the mass and composition of a main-sequence star determine its luminosity, temperature, and size.** The mass and composition of a main-sequence star determine its original position on the H-R diagram. The main sequence on the H-R diagram is actually a sequence of masses. This position connects its other properties such as its luminosity, temperature, and size.

? UNANSWERED QUESTIONS

- What is the upper limit for stellar mass? Both theory and observation have shown that there is a lower limit for stellar mass, approximately 0.08 M_{Sun} with temperature about 2000 K. However, neither theory nor observation has provided a definitive value for the upper limit. Many astronomers believe that the upper limit lies somewhere around 150–200 M_{Sun}. The very first stars that formed in the universe might have been even larger, but they are no longer around to observe.
- Are there likely to be planets with life orbiting around main-sequence stars of type M? These low-luminosity stars are the most common type of stars in the Milky Way, and the Kepler telescope has detected many planets orbiting these stars. However, the habitable zone of an M star is very close to the star, so that the planet may be strongly affected by streams of charged particles blowing off of the star. This radiation may decrease or completely strip a planet of its atmosphere, unless it has a strong protective magnetic field. Another complication is that a close-in planet is likely to be tidally locked to the star, so that one hemisphere of the planet receives light and the other hemisphere is permanently dark. This imbalance of light and heat might make the planet uninhabitable.

Questions and Problems

Test Your Understanding

1. Star A and star B are nearly the same distance from Earth. Star A is half as bright as star B. Which of the following statements must be true?
   a. Star B is farther away than star A.
   b. Star B is twice as luminous as star A.
   c. Star B is hotter than star A.
   d. Star B is larger than star A.

2. Star A and star B are two nearby stars. Star A is blue, and star B is red. Which of the following statements must be true?
   a. Star A is hotter than star B.
   b. Star A is cooler than star B.
   c. Star A is farther away than star B.
   d. Star A is more luminous than star B.

3. Star A and star B are two stars nearly the same distance from Earth. Star A is blue, and star B is red, but they have equal brightness. Which of the following statements is true?
   a. Star A is more luminous than star B.
   b. Star A is larger than star B.
   c. Star A is smaller than star B.
   d. Star A is less luminous than star B.

4. If a star has very weak hydrogen lines and is blue, what does that most likely mean?
   a. The star is too hot for hydrogen lines to form.
   b. The star has no hydrogen.
   c. The star is too cold for hydrogen lines to form.
   d. The star is moving too fast to measure the lines.

5. Star A and star B are a binary system. The Doppler shift of star A's absorption lines is 3 times the Doppler shift of star B's absorption lines. Which of the following statements is true?
   a. Star A is 3 times as massive as star B.
   b. Star A is one-third as massive as star B.
   c. Star A is closer than star B.
   d. The binary pair is moving toward Earth, but star A is farther away.

6. Star A and star B are two red stars at nearly the same distance from Earth. Star A is many times brighter than star B. Which of the following statements is true?
   a. Star A is a main-sequence star, and star B is a red giant.
   b. Star A is a red giant, and star B is a main-sequence star.
   c. Star A is hotter than star B.
   d. Star A is a white dwarf, and star B is a red giant.

7. Star A and star B are two blue stars at nearly the same distance from Earth. Star A is many times brighter than star B. Which of the following statements is true?
   a. Star A is a main-sequence star, and star B is a red giant.
   b. Star A is a main-sequence star, and star B is a blue giant.
   c. Star A is a white dwarf, and star B is a blue giant.
   d. Star A is a blue giant, and star B is a white dwarf.

8. In which region of an H-R diagram would you find the main-sequence stars with the widest habitable zones?
   a. upper left
   b. upper right
   c. center
   d. lower left
   e. lower right

9. Star A is more massive than star B. Both are main-sequence stars. Therefore, star A is __________ than star B. (Choose all that apply.)
   a. more luminous
   b. less luminous
   c. hotter
   d. colder
   e. larger
   f. smaller

10. A telescope on Mars would be able to measure the distances to more stars than can be measured from Earth because
    a. the resolution of the telescope would be better.
    b. Mars has a thin atmosphere.
    c. it would be closer to the stars.
    d. the parallax "baseline" would be longer.

11. Star A and star B are two nearby stars. Star A has a parallactic angle 4 times as large as star B's. Which of the following statements is true?
    a. Star A is one-quarter as far away as star B.
    b. Star A is 4 times as far away as star B.
    c. Star A has moved through space one-quarter as far as star B.
    d. Star A has moved through space 4 times as far as star B.

12. Star A appears twice as bright as star B, but is also twice as far away. Star A is __________ as luminous as star B.
    a. 8 times
    b. 4 times
    c. twice
    d. half

13. Table 13.1 shows two ways of reporting the amount of an element in the Sun. The percentage of hydrogen drops when changing from percentage by number of atoms to percentage by mass. But the percentage of helium grows. Why?
    a. Hydrogen is more massive than helium.
    b. Helium is more massive than hydrogen.
    c. Hydrogen is located in a different part of the Sun.
    d. It is difficult to measure the mass of hydrogen.

14. Capella (in the constellation Auriga) is the sixth brightest star in the sky. When viewed with a high-power telescope, it is clear that Capella is actually two pairs of binary stars: the first pair are G-type giants; the second pair are M-type main-sequence stars. What color does Capella appear to be?
    a. red
    b. yellow
    c. blue
    d. color cannot be determined from this information

15. An eclipsing binary system has a primary eclipse (star A is eclipsed by star B) that is deeper (more light is removed from the light curve) than the secondary eclipse (star B is eclipsed by star A). What does this tell you about stars A and B?
    a. Star A is hotter than star B.
    b. Star B is hotter than star A.
    c. Star B is larger than star A.
    d. Star B is moving faster than star A.

Thinking about the Concepts

16. The distances of nearby stars are determined by their parallaxes. Why is there greater uncertainty in the distances of stars that are farther from Earth?

17. To know certain properties of a star, you must first determine the star's distance. For other properties, knowledge of distance is not necessary. Explain why an astronomer does or does not need to know a star's distance to determine each of the following properties: size, mass, temperature, color, spectral type, and chemical composition? In each case, state your reason(s).

18. Albireo, in the constellation Cygnus, is a visual binary system whose two components can easily be seen with even a small, amateur telescope. Viewers describe the brighter star as "golden" and the fainter one as "sapphire blue."
    a. What does this description tell you about the relative temperatures of the two stars?
    b. What does this description tell you about their respective sizes?

19. Very cool stars have temperatures around 2500 K and emit Planck spectra with peak wavelengths in the red part of the spectrum. Do these stars emit any blue light? Explain your answer.

20. The stars Betelgeuse and Rigel are both in the constellation Orion. Betelgeuse appears red, and Rigel is bluish white. To the eye, the two stars seem equally bright. If you can compare the temperature, luminosity, or size from just this information, do so. If not, explain why.

21. Explain why the stellar spectral types (O, B, A, F, G, K, M) are not in alphabetical order. What sequence of temperatures is defined by these spectral types?

22. Other than the Sun, the only stars whose mass astronomers can measure *directly* are those in eclipsing or visual binary systems. Why? How do astronomers estimate the masses of stars that are not in eclipsing or visual binary systems?

23. Once the mass of a certain spectral type of star located in a binary system has been determined, it can be assumed that all other stars of the same spectral type and luminosity class have the same mass. Why is this a reasonable assumption?

24. Explain why the Kepler Mission is finding eclipsing binary stars while it is searching for extrasolar planets using the transit method.

25. Scientific advances often require the participation of scientists from all over the world, working on the same problem over many decades, even centuries. Compare and contrast this mode of "collaboration" with collaborations in your courses (perhaps on final projects or papers). What mechanisms must be in place to allow scientists to collaborate across space and time in this way?

26. What would happen to our ability to measure stellar parallax if we were on the planet Mars? What if we were on Venus or Jupiter?

27. In Figure 13.7, there is an absorption line at about 410 nm that is weak for O stars and weak for G stars but very strong in A stars. This particular line comes from the transition from the second excited state of hydrogen up to the sixth excited state. Why is this line weak in O stars? Why is it weak in G stars? Why is it strongest in the middle of the range of spectral types?

28. Which kinds of binary systems are best observed edge-on? Which kind are best observed face-on?

29. In Figure 13.10, two stars orbit a common center of mass.
    a. Explain why star 2 has a smaller orbit than star 1.
    b. Re-sketch this picture for the case where star 1 has a very low mass, perhaps close to that of a planet.
    c. Re-sketch this picture for the case where star 1 and star 2 have the same mass.

30. If our Sun were a blue main-sequence star, and Earth was still 1 AU from the Sun, would you expect Earth to be in the habitable zone? What about if our Sun were a red main-sequence star?

Applying the Concepts

31. Look at Figure 13.1b. Suppose the figure included a third star, located 4 times as far away as star A. How much less than star A would it appear to move each year? How much less than star B?

32. Suppose you see an object jump from side to side by half a degree as you blink back and forth between your eyes. How much farther away is an object that moves only one-third of a degree?

33. Logarithmic (log) plots show major steps along an axis scaled to represent equal factors, most often factors of 10. Why do astronomers sometimes use a log plot instead of the more conventional linear plot? Is the horizontal axis of the H-R diagram in Figure 13.15 logarithmic or linear?

34. Examine Figure 13.5. This figure is plotted logarithmically on both axes. The luminosities are in units of solar luminosities.
    a. How much more luminous than the Sun is a star on the far right side of the plot?
    b. How much less luminous than the Sun is a star on the far left side of the plot?

35. Study Figure 13.17. Compared to the Sun, how luminous, large, and hot is a star that has 10 times the mass of the Sun?

36. Sirius, the brightest star in the sky, has a parallax of 0.379 arcsec. What is its distance in parsecs? in light-years? How long does it take light to reach Earth?

37. Sirius is actually a binary pair of two A-type stars. The brighter of the two stars is called the "Dog Star" and the fainter is called the "Pup Star" because Sirius is in the constellation Canis Major (meaning "Big Dog"). The Dog Star appears about 6,800 times brighter than the Pup Star, even though both stars are at the same distance from Earth. Compare the temperatures, luminosities, and sizes of these two stars.

38. Sirius and its companion orbit around a common center of mass with a period of 50 years. The mass of Sirius is 2 times the mass of the Sun.
    a. If the orbital velocity of the companion is 2.35 times greater than that of Sirius, what is the mass of the companion?
    b. What is the semimajor axis of the orbit?

39. Sirius is 25 times more luminous than the Sun, and Polaris (the "North Pole Star") is 2,500 times more luminous than the Sun. Sirius appears 24 times brighter than Polaris. How much farther away is Polaris than Sirius? Use your answer from problem 36 to find the distance of Polaris in light-years.

40. Betelgeuse (in Orion) has a parallax of 0.00763 ± 0.00164 arcsec, as measured by the Hipparcos satellite. What is the distance to Betelgeuse, and what is the uncertainty in that measurement?

41. Rigel (also in Orion) has a Hipparcos parallax of 0.00412 arcsec. Given that Betelgeuse and Rigel appear equally bright in the sky, which star is actually more luminous. Knowing that Betelgeuse appears reddish while Rigel appears bluish white, which star would you say is larger and why?

42. The Sun is about 16 trillion (1.6×10^{13}) times brighter than the faintest stars visible to the naked eye.
    a. How far away (in astronomical units) would an identical solar-type star be if it were just barely visible to the naked eye?
    b. What would be its distance in light-years?

43. Study Figure 13.9. If $m_1 = m_2$, where would the center of mass be located? If $m_1 = 2m_2$, where would the center of mass be located?

44. Find the peak wavelength of blackbody emission for a star with a temperature of about 10,000 K. In what region of the spectrum does this wavelength fall? What color is this star?

45. About 1,470 watts (W) of solar energy hits each square meter of Earth's surface. Use this value and the distance to the Sun to calculate the Sun's luminosity.

USING THE WEB

46. Go to the European Space Agency's Gaia mission website (http://esa.int/science/gaia). How will it help astronomers determine the distances to more stars? Why is it better to make parallax measurements from space than from the ground on Earth? Have any data been released?

47. Go to the "Eclipsing Binary Stars Lab" website (http://astro.unl.edu/naap/ebs/ebs.html). Click on "Eclipsing Binary Simulator." Select preset Example 1, in which the two stars are identical. The animation will run with inclination 90° and show a 50 percent eclipse. What happens when you slowly change your viewing angle to the system—the inclination. How does this change the eclipse? At what value of inclination do you no longer see eclipses? What does the system look like at 0°? Reset the inclination to 90° and adjust the separation of the two stars. How does the light curve change when the separation is larger or smaller? Now make the two stars different. Change star 2 so that its radius is 3.0 R_{Sun} and its temperature is 4000 K. At what value of inclination do you no longer see eclipses? What types of eclipsing binary systems do you think are the easiest to detect?

48. Go to the Kepler home page (http://kepler.nasa.gov) and mouse over "Confirmed Planets" on the upper right. How many eclipsing binary stars has Kepler found? Go to the Kepler Eclipsing Binary Catalog (http://keplerebs.villanova.edu) to see what new observations look like. Pick a few stars to study. What is the inclination ("sin i")? Look at the last 2 columns ("Figures"). The "raw" and "dtr" figures are rough, but the "pf" figure shows a familiar light curve. How deep is the eclipse; that is, how much lower is the "normalized flux" during maximum eclipse?

49. Do a search for a photograph of your favorite constellation (or go outside and take a picture yourself). Can you see different colors in the stars? What do the colors tell you about the surface temperatures of the stars? From your photograph, can you tell which are the three brightest stars in the constellation? These stars will be named "alpha" (α), "beta" (β), and "gamma" (γ) for that constellation. Look up the constellation online and see if you chose the right stars. What are their temperatures and luminosities? What are their distances?

50. Citizen science: Go to the website for the Stellar Classification Online Public Exploration (SCOPE) project (http://scope.pari.edu/takepart.php). This project uses crowdsourcing to classify stars seen on old photographic plates of photographs taken in the Southern Hemisphere. Create an account, review the science and the FAQ, and then click on "To Take Part" to see some practice examples. Then go to "Classify," choose a photographic plate, and classify a few stars.

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

EXPLORATION

The H-R Diagram

digital.wwnorton.com/astro5

Open the "HR Explorer" interactive simulation for this chapter at the Student Site at the Digital Landing Page. This simulation enables you to compare stars on the H-R diagram in two ways. You can compare an individual star (marked by a red *X*) to the Sun by varying its properties in the box in the left half of the window. Or you can compare groups of the nearest and brightest stars. Play around with the controls for a few minutes to familiarize yourself with the simulation.

Begin by exploring how changes to the properties of the individual star change its location on the H-R diagram. First, press the "Reset" button at the top right of the window.

Decrease the temperature of the star by dragging the temperature slider to the left. Notice that the luminosity remains the same. Because the temperature has decreased, each square meter of star surface must be emitting less light. What other property of the star changes in order to keep the total luminosity of the star constant?

Predict what will happen when you slide the temperature slider all the way to the right. Now do it. Did the star behave as you expected?

1 As you move to the left across the H-R diagram, what happens to the radius?

____

____

2 What happens as you move to the right?

____

____

Press "Reset" and experiment with the luminosity slider.

3 As you move up on the H-R diagram, what happens to the radius?

____

____

4 What happens as you move down?

____

____

Press "Reset" again and predict how you would have to move the slider bars to move your star into the red giant portion of the H-R diagram (upper right). Adjust the slider bars until the star is in that area. Were you correct?

5 How would you adjust the slider bars to move the star into the white dwarf area of the H-R diagram?

____

____

Press the "Reset" button and explore the right-hand side of the window. Add the nearest stars to the graph by clicking their radio button under "Plotted Stars." Using what you learned above, compare the temperatures and luminosities of these stars to the Sun (marked by the *X*).

6 Are the nearest stars generally hotter or cooler than the Sun?

____

7 Are the nearest stars generally more or less luminous than the Sun?

____

Press the radio button for the brightest stars. This action will remove the nearest stars and add the brightest stars in the sky to the plot. Compare these stars to the Sun.

8 Are the brightest stars generally hotter or cooler than the Sun?

____

9 Are the brightest stars generally more or less luminous than the Sun?

____

10 How do the temperatures and luminosities of the brightest stars in the sky compare to the temperatures and luminosities of the nearest stars? Does this information support the claim in the chapter that there are more low-luminosity stars than high-luminosity stars? Explain.

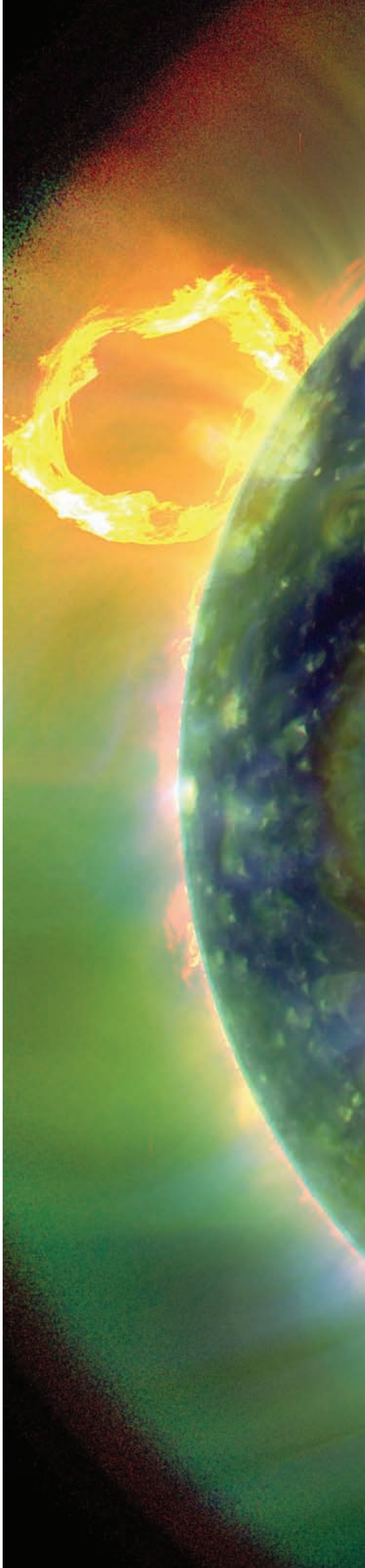
____

____

____

____

14 Our Star—The Sun



Because the Sun is the only star close to Earth, much of the detailed knowledge about stars has come from studying the Sun. In Chapter 13, we looked at the physical properties of distant stars, including their mass, luminosity, size, temperature, and chemical composition. In this chapter, we ask fundamental questions about Earth's local star. How does the Sun work? Where does it get its energy? How has its luminosity been able to remain so constant over the billions of years since the Solar System formed?

LEARNING GOALS

By the conclusion of this chapter, you should be able to:

LG 1 Describe the balance between the forces that determine the structure of the Sun.

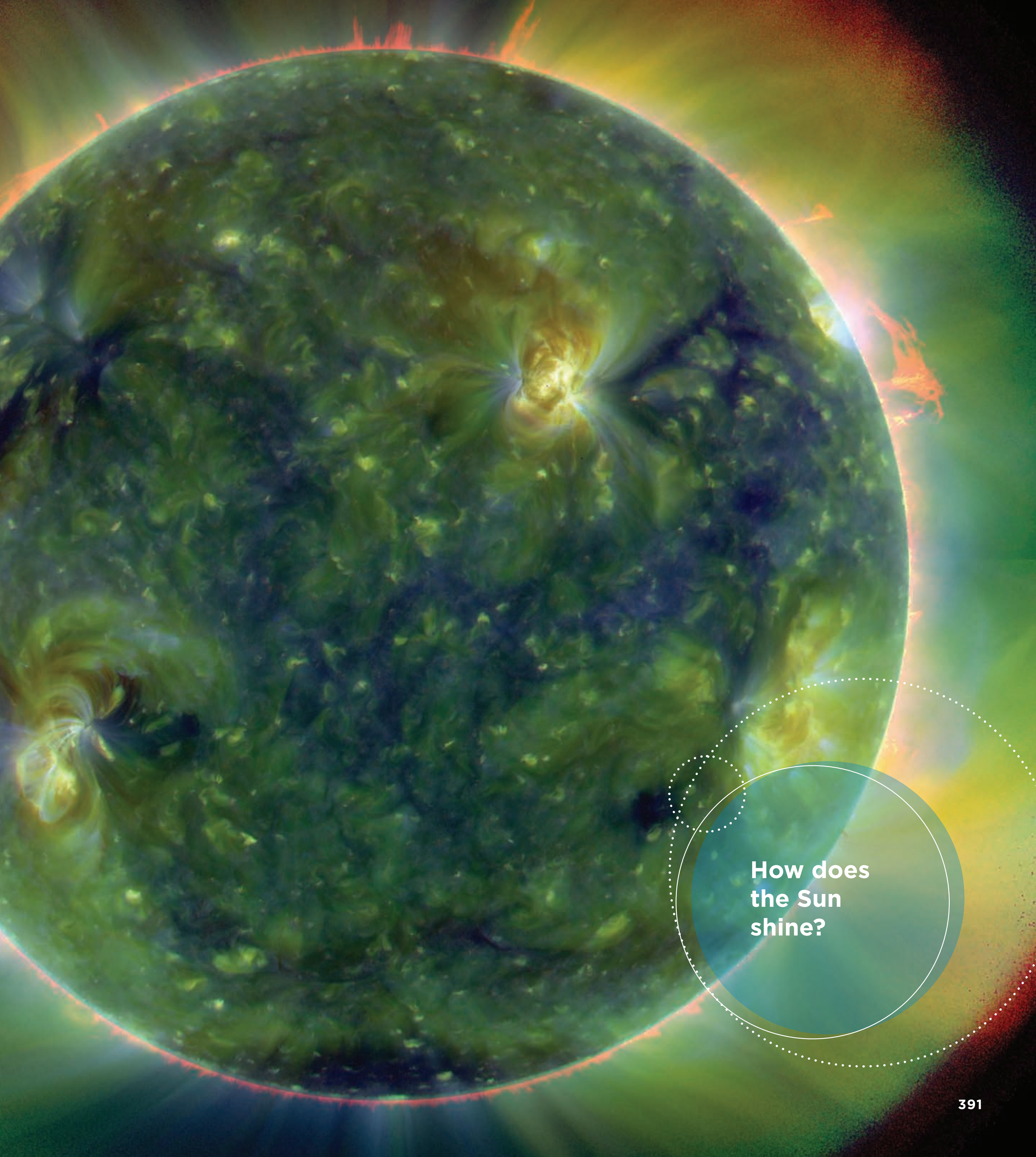
LG 2 Explain how mass is converted into energy in the Sun's core and how long it will take the Sun to use up its fuel.

LG 3 Sketch a physical model of the Sun's interior, and list the different ways that energy moves outward from the Sun's core toward its surface.

LG 4 Describe how observations of solar neutrinos and seismic vibrations on the surface of the Sun test astronomers' models of the Sun.

LG 5 Describe the solar activity cycles of 11 and 22 years, and explain how these cycles are related to the Sun's changing magnetic field.

This image is a combination of several extreme ultraviolet images of the Sun from the Solar Dynamics Observatory. ▶▶▶



How does the Sun shine?

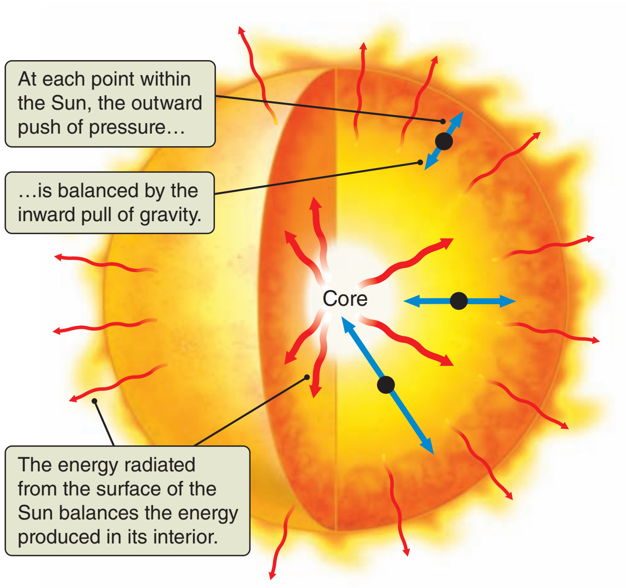


Figure 14.1 The structure of the Sun is determined by the balance between the forces of pressure and gravity and the balance between the energy generated in its core and the energy radiated from its surface.

14.1 The Sun Is Powered by Nuclear Fusion

Energy from the Sun is responsible for daylight, for Earth's weather and seasons, and for terrestrial life itself. At a luminosity of 3.85×10^{26} watts (W), the Sun produces more energy in a second than all of the electric power plants on Earth could generate in a half-million years. In this section, we will look at energy production in the Sun.

Hydrostatic Equilibrium

Geologists learn about the interior of Earth by using a combination of physics, detailed computer models, and experiments that test the predictions of those models. The task of exploring the interior of the Sun is much the same. Like Earth's structure, the structure of the Sun is governed by a number of physical processes and relationships. Using physics, chemistry, and the properties of matter and radiation, astronomers can express these processes and relationships as mathematical equations. They then solve these equations and create a model of the Sun. One of the great successes of 20th century astronomy was the construction of a physical model of the Sun that agrees with observations of the mass, composition, size, temperature, and luminosity of the real thing.

The structure of the Sun is a matter of balance between the pressure outward and the force of gravity inward: this balance is known as **hydrostatic equilibrium**. The pressure results from energy finding its way to the surface of the Sun from deep in its interior. To understand how hydrostatic equilibrium affects the Sun, we need to know how these forces are produced and how they continually change to balance each other.

The balance between the forces due to pressure and gravity is illustrated in **Figure 14.1**. The Sun is a huge ball of hot gas. Deep in the Sun's interior, the outer layers press downward because of gravity, producing a large inward force. To maintain balance, the outward force due to pressure must be equally large. If gravity were not balanced by pressure, the Sun would collapse. If pressure were not balanced by gravity, the Sun would blow itself apart. At every point within the Sun's interior, the pressure must be just enough to hold up the weight of all the layers above that point. If the Sun were not in a stable hydrostatic equilibrium, forces within it would not be in balance, and the size of the Sun would change.

Hydrostatic equilibrium becomes an even more powerful concept when combined with the way gases behave. Deeper in the interior of the Sun, the weight of the material above becomes greater, and hence the pressure must increase. In a gas, higher pressure means higher density and/or higher temperature. **Figure 14.2** shows how conditions vary inside the Sun. As illustrated in the graphs in Figure 14.2b, calculations show that toward the center of the Sun, the pressure, the density, and the temperature of the gas increase.

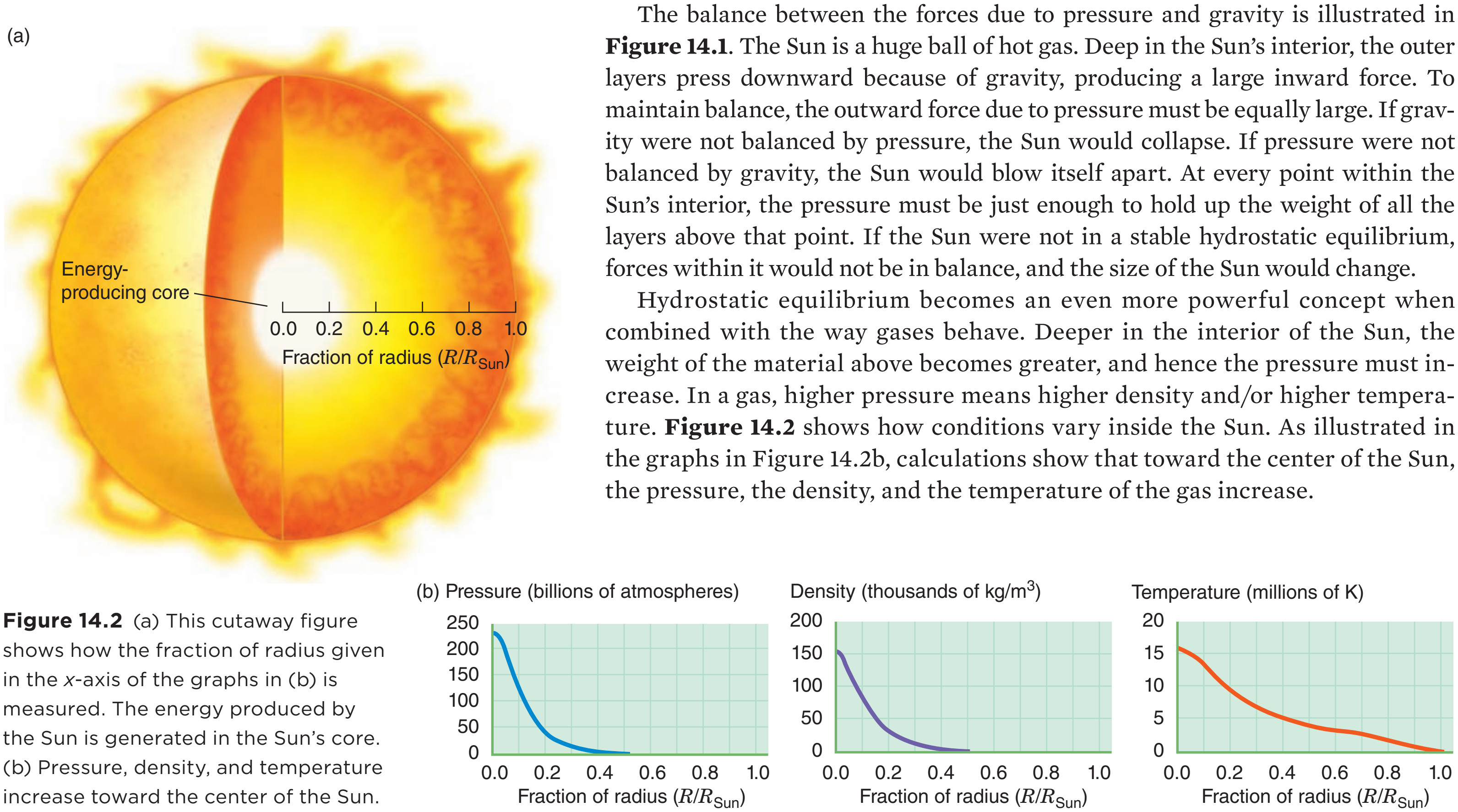


Figure 14.2 (a) This cutaway figure shows how the fraction of radius given in the *x*-axis of the graphs in (b) is measured. The energy produced by the Sun is generated in the Sun's core. (b) Pressure, density, and temperature increase toward the center of the Sun.

CHECK YOUR UNDERSTANDING 14.1

Hydrostatic equilibrium in the Sun means that: (a) the Sun does not change over time; (b) the Sun absorbs and emits equal amounts of energy; (c) pressure balances the weight of overlying layers; (d) energy produced in the core per unit time equals energy emitted at the surface per unit time.

Nuclear Fusion

A second fundamental balance within the Sun is the balance of energy (see Figure 14.1). Stars like the Sun are remarkably stable objects. Geological records show that the luminosity of the Sun has remained nearly constant for billions of years. To remain in balance, the Sun must produce just enough energy in its interior to replace the energy radiated away from its surface. This energy balance tells us how much energy must be produced in the interior of the Sun and how that energy finds its way from the interior to the Sun's surface, where it is radiated away. Models of stellar evolution indicate that the luminosity of the Sun is increasing with time, but very, very slowly. The Sun's luminosity 4.5 billion years ago was about 70 percent of its current luminosity.

The amount of energy produced by the Sun each second is truly astronomical: 3.85×10^{26} W. One of the most basic questions facing the pioneers of stellar astrophysics was how the Sun and other stars get their energy. In the 19th century, physicists proposed that the Sun was slowly shrinking, and that the core was heating up as a result of this gravitational contraction. However, calculations soon showed that this would power the Sun for only millions of years. Geological and biological evidence available at the time suggested that Earth was tens of millions or hundreds of millions of years old. By the early 20th century, radiometric dating suggested that Earth was more than a billion years old, and therefore gravitational contraction could not be the source of the Sun's energy. In the 1930s, using theoretical and laboratory physics, nuclear physicists concluded that the Sun's energy comes from nuclear reactions at its core, capable of powering the star for billions of years.

Recall from Chapter 5 that the nucleus of most hydrogen atoms consists of a single proton. Nuclei of all other atoms are built from a mixture of protons and neutrons. Most helium nuclei, for example, consist of two protons and two neutrons. Protons have a positive electric charge, and neutrons have no electric charge. Because like charges repel, and the closer they are the stronger the force, all of the protons in an atomic nucleus are continually repelling each other with a tremendous force. The nuclei of atoms should fly apart due to electric repulsion—yet atomic nuclei are held together by the **strong nuclear force**, which overcomes this repulsion. However, the strong nuclear force acts only over very short distances, of the order 10^{-15} meter, about the size of the atomic nucleus, or about a hundred-thousandth the size of an atom.

Compared to the energy required to free an electron from an atom, the amount of energy required to tear a nucleus apart is enormous. Conversely, when an atomic nucleus (with a mass up to and including the nucleus of iron) is formed from component parts, energy is released. **Nuclear fusion**—the process of combining two less massive atomic nuclei into a single more massive atomic nucleus—occurs when atomic nuclei are brought close enough together for the strong nuclear force to overcome the force of electric repulsion, as illustrated in **Figure 14.3**. Many kinds of nuclear fusion can occur in stars. In main-sequence stars like the Sun, the primary energy generation process is the fusion of hydrogen into helium—a process often called **hydrogen burning** (even though it has nothing to

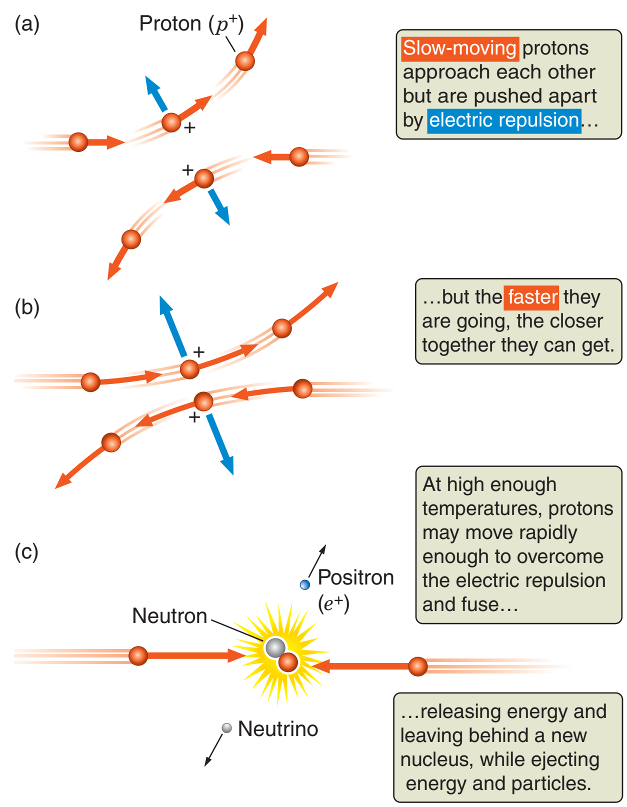


Figure 14.3 (a) Atomic nuclei are positively charged and electrically repel each other. (b) The faster that two nuclei are moving toward each other, the closer they will get before veering away. (c) At the temperatures and densities found in the centers of stars, nuclei can overcome this electric repulsion, so fusion takes place.

do with fire or other chemical combustion). The fusion of hydrogen into helium always takes several steps, but the net result is that four hydrogen nuclei become the one helium nucleus plus energy.

The energy produced in nuclear reactions comes from the conversion of mass into energy. The exchange rate between mass and energy is given by Einstein's famous equation, $E = mc^2$, in which E is energy, m is mass, and c^2 is the speed of light squared. For any nuclear reaction, we can determine the mass that is turned into energy by calculating the mass that is lost. To find this lost mass, we subtract the mass of the outputs from the mass of the inputs. In hydrogen burning, the inputs are four hydrogen nuclei, and the output is a helium nucleus plus energy. The mass of four separate hydrogen nuclei is 1.007 times greater than the mass of a single helium nucleus; so when hydrogen fuses to make helium, 0.7 percent of the mass of the hydrogen is converted to energy.

Although each fusion reaction produces a small amount of energy, the total mass of the Sun is very large, so there is much hydrogen to "burn." When the amount of energy produced by nuclear burning is compared with the luminosity of the Sun, we see that these reactions can power the Sun for 10 billion years, a time frame that is longer than the 4.6-billion-year age of the Solar System measured from radioactive dating. Details of this calculation are provided in **Working It Out 14.1**.

14.1 Working It Out The Source of the Sun's Energy

Like all stars, the Sun's lifetime is limited by the amount of fuel available to it. We can calculate how long the Sun will live by comparing the mass involved in nuclear fusion with the amount of mass available. Converting four hydrogen nuclei (protons) into a single helium nucleus results in a loss of mass. The mass of a single hydrogen nucleus is 1.6726×10^{-27} kilogram (kg). So, four hydrogen nuclei have a mass of 4 times that, or 6.6904×10^{-27} kg. The mass of a helium nucleus is 6.6447×10^{-27} kg, which is less than the mass of the four hydrogen nuclei. The amount of mass lost, m, is

$$m = 6.6904 \times 10^{-27}\,\text{kg} - 6.6447 \times 10^{-27}\,\text{kg} = 0.0457 \times 10^{-27}\,\text{kg}$$

We can write this as 4.57×10^{-29} kg—a mass loss of about 0.7 percent. Conversion of 0.7 percent of the mass of the hydrogen into energy might not seem very efficient—until we compare it with other sources of energy and discover that it is millions of times more efficient than even the most efficient chemical reactions.

Using Einstein's equation $E = mc^2$, where c is the speed of light (3×10^8 m/s), along with the definition of a joule ($1\text{ J} = 1\text{ kg m}^2/\text{s}^2$), we can calculate the energy released by this mass-to-energy conversion:

$$E = mc^2 = (4.57 \times 10^{-29}\,\text{kg}) \times (3.00 \times 10^8\,\text{m/s})^2 = 4.11 \times 10^{-12}\,\text{J}$$

Each reaction that takes four hydrogen nuclei and turns them into a helium nucleus releases 4.11×10^{-12} J of energy, which doesn't seem like very much. But atoms are very small. Fusing a single kilogram of hydrogen into helium releases about 6.3×10^{14} J of energy—about the equivalent of the chemical energy released in burning 100,000 barrels of oil. To see how much the Sun must be fusing per second to produce its current luminosity, we divide the luminosity of the Sun by this amount of energy per kilogram:

$$\frac{\text{Luminosity of Sun}}{\text{Energy per kilogram}} = \frac{4 \times 10^{26}\,\text{J/s}}{6.3 \times 10^{14}\,\text{J/kg}} = 6.2 \times 10^{11}\,\text{kg/s}$$

For the Sun to produce as much energy as it does, it must convert roughly 620 billion kg of hydrogen into helium every second (and about 4 billion kg of matter is converted to energy in the process). The Sun has been burning hydrogen at this rate for at least the age of Earth and the Solar System—4.6 billion years. How much longer will the Sun last?

Astronomers estimate that only 10 percent of the Sun's total mass will ever be involved in fusion, because the other 90 percent will never get hot enough or dense enough for the strong nuclear force to make fusion happen. Ten percent of the mass of the Sun is $(0.1) \times (2 \times 10^{30})$ kg, or 2×10^{29} kg. That is the amount of "fuel" the Sun has available. The Sun consumes hydrogen at a rate of 620 billion kg/s, so each year the Sun consumes:

$$M_{\text{year}} = (6.2 \times 10^{11}\,\text{kg/s}) \times (3.16 \times 10^7\,\text{s/yr}) = 2 \times 10^{19}\,\text{kg/yr}$$

If we know how much fuel the Sun has (2×10^{29} kg), and we know how much the Sun burns each year (2×10^{19} kg/yr), then we can divide the amount by the rate to find the lifetime of the Sun:

$$\text{Lifetime} = \frac{M_{\text{fuel}}}{M_{\text{year}}} = \frac{2 \times 10^{29}\,\text{kg}}{2 \times 10^{19}\,\text{kg/yr}} = 10^{10}\,\text{yr}$$

When the Sun was formed, it had enough fuel to power it for about 10 billion years. The Sun is nearly halfway through its lifetime of hydrogen burning.

Energy is produced in the Sun's innermost region, the **core**, where the conditions are the most extreme. The density of matter in the core is about 150 times the density of water, and the temperature is about 15 million kelvins (K). Under these conditions, the atomic nuclei have tens of thousands of times more kinetic energy than that of atoms at room temperature and can slam into each other hard enough to overcome the electric repulsion, allowing the strong nuclear force to act (Figure 14.3c). In hotter and denser gases, such collisions happen more frequently. For this reason, the rate of nuclear fusion reactions is extremely sensitive to the temperature and the density of the gas, which is why these energy-producing collisions are concentrated in the Sun's core. Half of the energy produced by the Sun is generated within the inner 9 percent of the Sun's radius, or less than 0.1 percent of the volume of the Sun.

▶II **AstroTour:** The Solar Core

The conversion of four hydrogen nuclei to one helium nucleus is the most significant source of energy in main-sequence stars. Hydrogen is the most abundant element in the universe, so it is the most abundant source of nuclear fuel at the beginning of a star's lifetime. Hydrogen is also the easiest type of atom to fuse. Hydrogen nuclei—protons—have an electric charge of +1. The electric barrier that must be overcome to fuse protons is the repulsion of one proton against another. To fuse 2 carbon nuclei together, for example, the repulsion of the six protons in one carbon nucleus pushing against the six protons in another carbon nucleus must be overcome. The repulsion between two carbon nuclei is 36 times stronger than that between two hydrogen nuclei. Therefore, hydrogen fusion occurs at a much lower temperature than any other type of nuclear fusion.

The Proton-Proton Chain

To test the theory that the Sun shines because of nuclear fusion, astronomers can analyze the predicted by-products of the nuclear reactions. In the core of the Sun and in other low-mass stars, hydrogen burning takes place in a series of nuclear reactions called the **proton-proton chain**, which has three different branches. The most important branch, responsible for about 85 percent of the energy generated in the Sun, consists of three steps, illustrated in **Figure 14.4**. Each step produces particles and/or energy in the form of light. We will begin by following the creation of the helium nucleus, and then go back to find out what happens to the other products of the reaction.

Follow along in Figure 14.4 as we step through the proton-proton chain. The nucleus of hydrogen consists of one proton. In the first step, two protons fuse. During this process, one of the protons is transformed into a neutron. To conserve spin and charge, two particles are emitted: a positively charged particle called a **positron** and a neutral particle called a **neutrino**. Energy is also emitted in the form of photons carrying electromagnetic radiation. The new atomic nucleus formed by the first

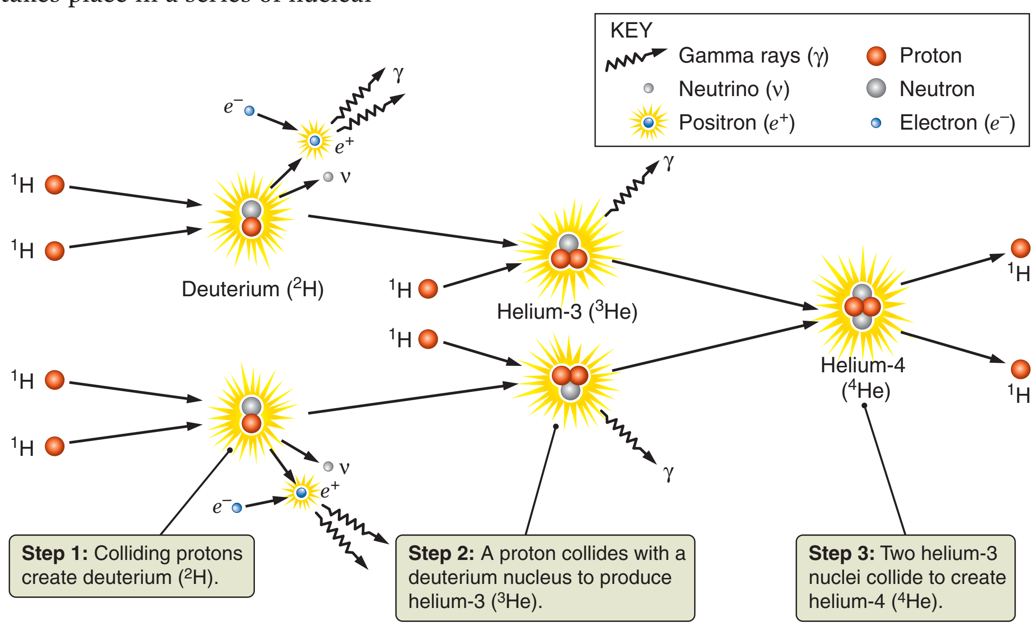


Figure 14.4 The Sun and all other main-sequence stars get their energy by fusing the nuclei of four hydrogen atoms together to make a single helium atom. In the Sun, about 85 percent of the energy produced comes from the branch of the proton-proton chain illustrated here.

step in the chain consists of a proton and a neutron. Recall from Chapter 5 that an *isotope* of an element has the same number of protons and a different number of neutrons. Thus, the new atomic nucleus is still hydrogen because it has only one proton. This particular isotope of hydrogen is common enough that it has its own name—deuterium, written as ^{2}H (H is the element symbol for hydrogen, which always has one proton, and 2 is the *atomic mass number*—the total number of protons and neutrons in the nucleus).

In the second step of the proton-proton chain, another proton slams into the deuterium nucleus, forming the nucleus of an isotope of helium, ^{3}He, consisting of two protons and a neutron. The energy released in this step is carried away as a highly energetic gamma-ray photon. Notice that the first two steps are shown twice, along the top of the figure and the bottom, because these steps must occur twice to produce a single ^{4}He nucleus.

In the third and final step of the proton-proton chain, two ^{3}He nuclei collide and fuse, producing an ordinary ^{4}He nucleus and ejecting two protons in the process. The energy released in this step is the kinetic energy of the helium nucleus and two ejected protons. Overall, four hydrogen nuclei have combined to form one helium nucleus.

Now let's go back and look at what happens to the other products of the reaction. In step 1, a positron—a particle of antimatter—is produced. **Antimatter** particles have the same mass as a corresponding matter particle but have opposite values of other properties, such as charge. The positron (e^+) is the antimatter counterpart of an electron (e^-). When matter (electrons) and antimatter (positrons) meet, they annihilate each other, and their total mass is converted to energy in the form of gamma-ray photons (γ). This happens to the emitted positrons inside the Sun, and the emitted photons from the annihilation carry away part of the energy released when the two protons fused. These photons heat the surrounding gas. The gamma rays emitted in step 2 similarly heat the gas. The thermal energy produced in the core of the Sun takes 100,000 years to find its way to the Sun's surface, and so the light we see from the Sun indicates what the Sun was doing 100,000 years ago.

The neutrino emitted in step 1 has a very different fate. Neutrinos are particles that have no charge, very little mass, and travel at nearly the speed of light. They interact weakly with ordinary matter, so weakly that the neutrino escapes from the Sun without further interactions with any other particles. The core of the Sun lies buried beneath 700,000 kilometers (km) of dense, hot matter, yet the Sun is transparent to neutrinos—essentially all of them travel into space as if the outer layers of the Sun were not there. Because they travel at nearly the speed of light, neutrinos from the center of the Sun arrive at Earth after only $8\frac{1}{3}$ minutes. Therefore, we can use them to probe what the Sun is doing today.

This dominant branch of the proton-proton chain can be written symbolically as follows:

Step 1: $^1\text{H} + {}^1\text{H} \rightarrow {}^2\text{H} + e^+ + \nu$ and then $e^+ + e^- \rightarrow \gamma + \gamma$

Step 2: $^2\text{H} + {}^1\text{H} \rightarrow {}^3\text{He} + \gamma$

Step 3: $^3\text{He} + {}^3\text{He} \rightarrow {}^4\text{He} + {}^1\text{H} + {}^1\text{H}$

Nebraska Simulation: Proton-Proton Animation

The rate of the proton-proton chain reaction depends on both temperature and density. At the temperature and pressure that exist within the Sun's core, the reaction rate is relatively slow—in fact, extremely slow compared to a nuclear bomb explosion. The Sun's slow nuclear fusion is fortunate for life on Earth: if its hydrogen burned quickly, the Sun would have exhausted its supply long ago, and life might not have had time to evolve.

The other 15 percent of the Sun's energy is generated by variations of the proton-proton chain. The most common variation happens in step 3, where ^{3}He fuses with an existing ^{4}He to create beryllium (^{7}Be), which decays to lithium (^{7}Li) and energy, and then the ^{7}Li plus one ^{1}H become two ^{4}He. In a less common variation, the beryllium combines with hydrogen to become boron (^{8}B), which then decays to beryllium and then to two ^{4}He. In both of these variations, ultimately four hydrogen nuclei become one helium nucleus.

CHECK YOUR UNDERSTANDING 14.2

When hydrogen is fused into helium, energy is released from: (a) gravitational collapse; (b) conversion of mass to energy; (c) the increase in pressure; (d) the decrease in the gravitational field.

14.2 Energy Is Transferred from the Interior of the Sun

Although geologists cannot travel deep inside Earth to find out how it is structured, they are able to build a model of its interior using data on how seismic waves travel during earthquakes. Similarly, astronomers can create a model of the Sun's interior using their knowledge of the balance of forces and energy within the Sun and an understanding of how energy moves from one place to another. These models can be tested by observations of waves traveling through the Sun and by studying neutrinos from the Sun.

Energy Transport

Some of the energy released by hydrogen burning in the core of the Sun escapes directly into space in the form of neutrinos. However, most of the energy heats the solar interior and then moves outward through the Sun to the surface, a process known as **energy transport**. Energy transport**,** a key determinant of the Sun's structure, can occur by *conduction*, *convection*, or *radiation*.

Conduction is important primarily in solids. For example, when you pick up a hot object, your fingers are heated by conduction. This happens because energetic thermal vibrations of atoms and molecules cause neighboring atoms and molecules to vibrate more rapidly as well. Conduction is typically ineffective in a gas because the atoms and molecules are too far apart to transmit vibrations to one another efficiently. Conduction does not play a key role in the transport of energy from the core of the Sun to its surface, but it will be relevant later when we discuss dying stars.

In the Sun, energy is transported by convection and radiation through different zones, as shown in **Figure 14.5**. The mechanism of energy transport from the center of the Sun outward depends on the decreasing temperature and density as the radius increases. First, energy moves outward through the inner layers of the

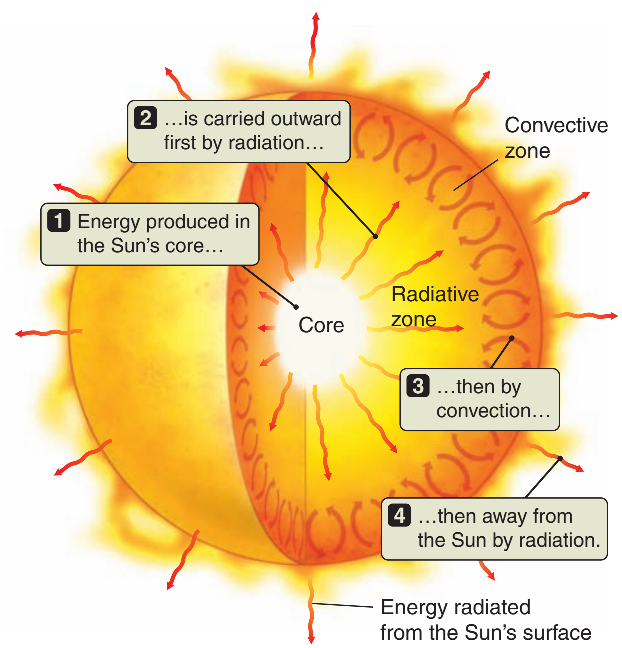


Figure 14.5 The interior structure of the Sun is divided into zones on the basis of where energy is produced and how it is transported outward.

Sun as radiation in the form of photons. Next, energy moves by convection in parcels of gas. Finally, energy radiates from the Sun's surface as light. We will look at each process in turn.

Near the core, **radiation** transfers energy from hotter to cooler regions via photons, which carry the energy with them. Consider a hotter region of the Sun located next to a cooler region, as shown in **Figure 14.6**. Recall from your study of radiation in Chapter 5 that the hotter region contains more (and more energetic) photons than the cooler region. More photons move from the hotter, very crowded region to the cooler, less crowded region than in the reverse direction. A net transfer of photons and photon energy occurs from the hotter region to the cooler region, and radiation carries energy outward from the Sun's core.

The transfer of energy from one point to another by radiation also depends on how freely radiation can move from one point to another within a star. The degree to which matter blocks the flow of photons through it is called **opacity**. The opacity of a material depends on many things, including the density of the material, its composition, its temperature, and the wavelength of the photons moving through it.

Astronomy in Action: Random Walk

Energy transfer by radiation is most efficient in regions with low opacity. The radiative zone (see Figure 14.5) is the region in the inner part of the Sun where the opacity is relatively low, and radiation carries the energy produced in the core outward through the star. This radiative zone extends about 70 percent of the way out toward the surface of the Sun. Even though this region's opacity is low enough for radiation to dominate convection as an energy transport mechanism, photons still travel only a short distance within the region before being absorbed, emitted, or deflected by matter, much like a beach ball being batted about by a

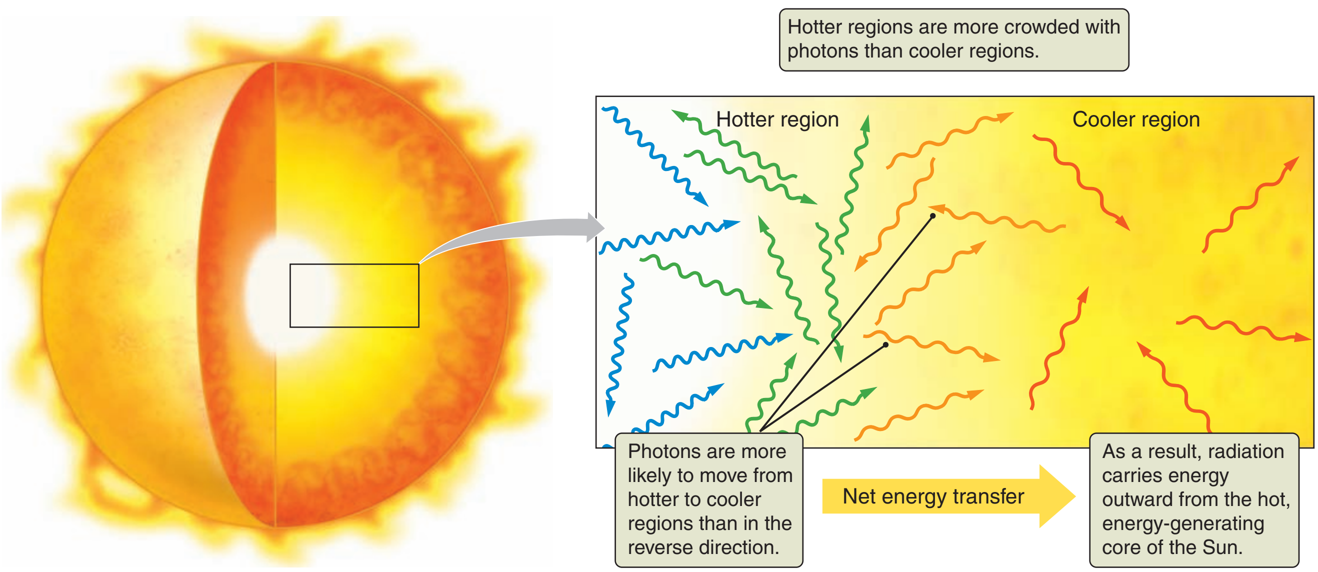


Figure 14.6 Higher-temperature regions deep within the Sun produce more radiation than do lower-temperature regions farther out. Although radiation flows in both directions, more radiation flows from the hotter regions to the cooler regions than from the cooler regions to the hotter regions. Therefore, radiation carries energy outward from the inner parts of the Sun. For simplicity, we have only included in this illustration the most common photons (those at the peak of the blackbody curve). Photons of all colors are present in all regions, and there are more of all kinds in the hotter regions and fewer of all kinds in the cooler regions.

crowd of people as illustrated in **Figure 14.7**. Each interaction sends the photon in an unpredictable direction—not necessarily toward the surface of the star. The distances between interactions are so short that, on average, it takes the energy of a gamma-ray photon produced in the interior of the Sun about 100,000 years to find its way to the outer layers of the Sun. Opacity holds energy within the interior of the Sun and lets it seep away only slowly. As it travels, the gamma-ray photon gradually becomes converted to lower-energy photons, emerging mostly as optical and infrared radiation from the surface.

From a peak of 15 million K at the center of the Sun, the temperature falls to about 100,000 K at the outer margin of the radiative zone. At this cooler temperature, the opacity is higher, so radiation is less efficient at carrying energy from one place to another. The energy that is flowing outward through the Sun "piles up" against this edge of the radiative zone.

Nearer the surface of the Sun, transfer by radiation becomes inefficient and the temperature changes quickly. Instead, **convection** takes over. Convection carries energy from the interior of a planet to its surface or from the Sun-heated surface of Earth upward through Earth's atmosphere. Convection also plays an important role in the transport of energy outward from the interior of the Sun. It transports energy by moving packets of hot gas, like hot-air balloons, which become buoyant and rise up through the lower-temperature gas above them, carrying energy with them. The solar convective zone (see Figure 14.5) extends from the outer boundary of the radiative zone outward to just below the visible surface of the Sun where evidence of convection can be seen in the bubbling surface (**Figure 14.8**).

In the outermost layers of the Sun, radiation again takes over as the primary mode of energy transport, and it is radiation that transports energy from the Sun's outermost layers off into space.

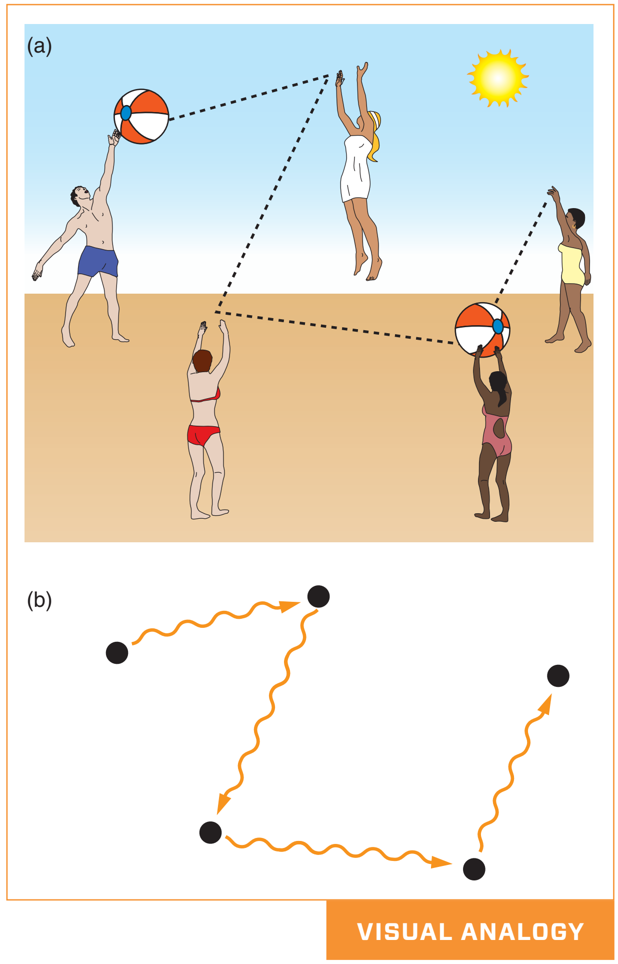


Figure 14.7 (a) When a crowd of people plays with a beach ball, the ball never travels very far before someone hits it, turning it in another direction. It often takes a ball a long time to make its way from one edge of the crowd to the other. (b) Similarly, when a photon travels through the Sun, it takes a long time for a photon to make its way out of the Sun.

Observing Neutrinos from the Core of the Sun

The model of energy production and energy transport in the Sun discussed above correctly matches observed global properties of the Sun such as its size, temperature, and luminosity. The nuclear fusion model of the Sun predicts exactly which nuclear reactions should be occurring in the core of the Sun and at what rate. The nuclear reactions that make up the proton-proton chain produce a vast number of neutrinos. Because neutrinos barely interact with other ordinary matter, almost all of the neutrinos produced in the heart of the Sun travel freely through the outer parts of the Sun and on into space as if the outer layers of the Sun were not there. Solar neutrinos produced in the core of the Sun, traveling at nearly the speed of light, take only $8\frac{1}{3}$ minutes to reach Earth, much quicker than the 100,000-year journey of photons.

Neutrinos interact so weakly with matter that they are extremely difficult to observe. However, the extremely large number of nuclear reactions in the Sun means that the Sun produces an enormous number of neutrinos. As you read this sentence, about 400 trillion solar neutrinos pass through your body. This happens even at night, as neutrinos easily pass through Earth. With this many neutrinos about, a neutrino detector does not have to detect a very large percentage of them to be effective.

A neutrino telescope looks very different from other telescopes. The first apparatus designed to detect solar neutrinos was built 1,500 meters underground,

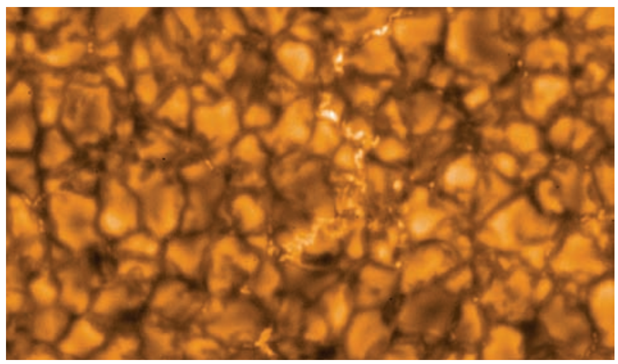


Figure 14.8 The top of the Sun's convective zone shows the bubbling of the surface caused by rising and falling packets of gas.

within the Homestake Mine in Lead, South Dakota. Astronomers filled a tank with 100,000 gallons of dry-cleaning fluid—C_2Cl_4, or perchloroethylene. Over the course of 2 days, astronomers predicted that roughly 10^{22} solar neutrinos passed through the Homestake detector. Of these, on average only one neutrino interacted with a chlorine atom within the fluid to form a radioactive isotope of argon. Over time, a measurable amount of argon was produced.

The Homestake experiment operated from the late 1960s to the early 1990s and detected this argon isotope—evidence of neutrinos from the Sun, confirming that nuclear fusion powers the Sun. However, astronomers noticed that they were measuring only one-third to one-half as many solar neutrinos as predicted by solar models. The difference between the predicted and measured number of solar neutrinos was called the **solar neutrino problem**.

One possible explanation of the solar neutrino problem was that the working model of the structure of the Sun was somehow wrong. This possibility seemed unlikely, however, because of the many other successful predictions of the solar model. A second possibility was that an understanding of the neutrino itself was incomplete. The neutrino was long thought to have zero mass, like photons, and to travel at the speed of light. But if neutrinos actually do have a tiny amount of mass, then particle physics suggests that solar neutrinos should oscillate—alternate back and forth among three different kinds of neutrinos: the electron, muon, and tau neutrinos. According to this explanation, early neutrino experiments could detect only the electron neutrino and consequently observed only about a third of the expected number of neutrinos. Since then, many other neutrino detectors have been built, each using different reactions to detect neutrinos of different energies or different types. Experiments at high-energy physics labs, nuclear reactors, and neutrino telescopes around the world have shown that neutrinos do have a nonzero mass and do oscillate among neutrino types.

Solving the solar neutrino problem is a good example of how science works—how a better model of the neutrino showed that the solar neutrino problem was real and not merely an experimental mistake, and how a single set of anomalous observations was later confirmed by other, more sophisticated experiments. All of this effort has led to a better understanding of basic physics (**Process of Science Figure**).

Probing the Sun's Interior

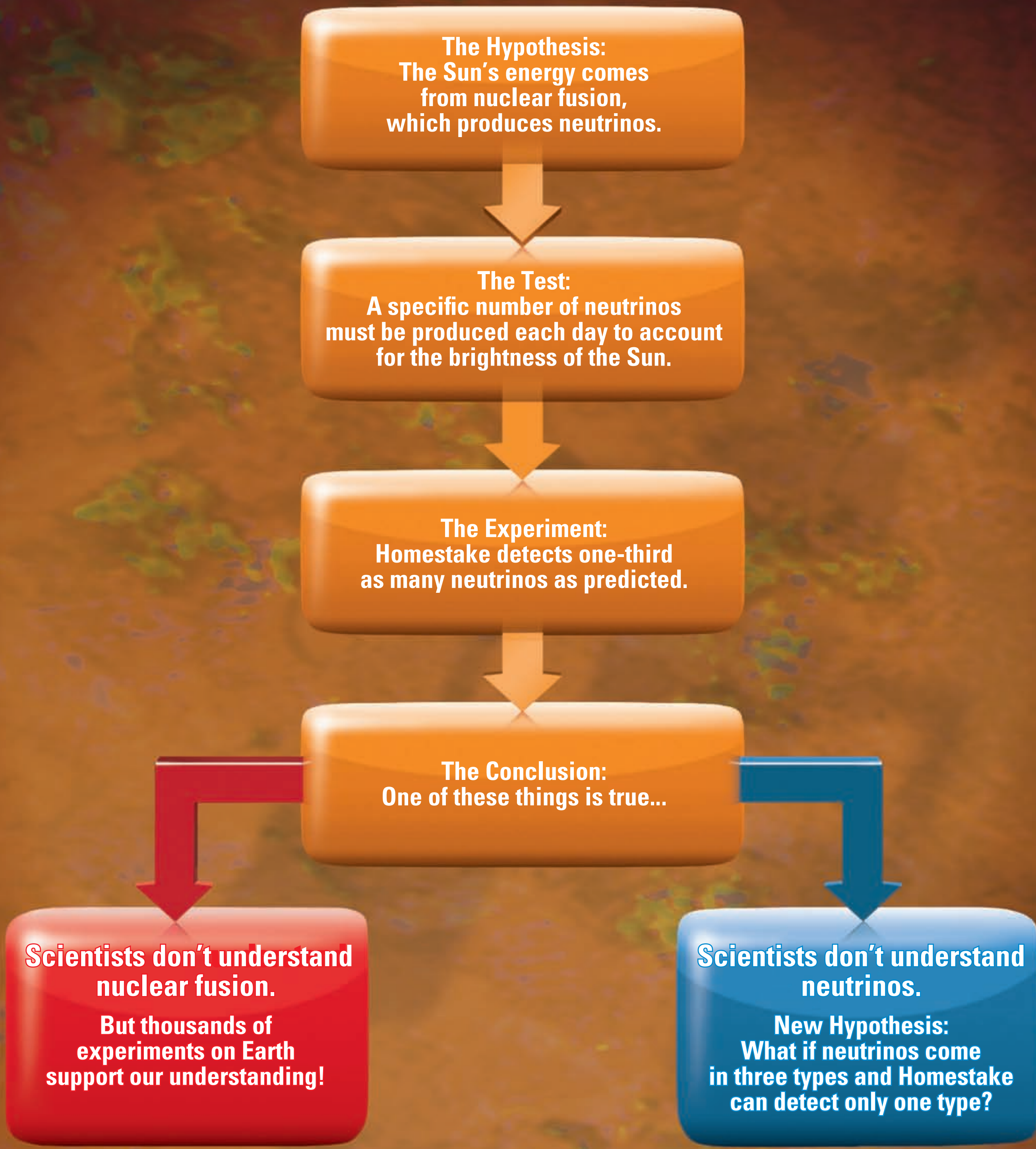
Models of Earth's interior predict how density and temperature change from place to place. These differences affect the seismic waves traveling through Earth, bending the paths that they travel. Geologists test models of Earth's interior by comparing measurements of seismic waves from earthquakes with model predictions of how seismic waves should travel through the planet.

Just as geologists use seismic waves from earthquakes to probe the interior of Earth, solar physicists use the surface oscillations of the Sun to probe the solar interior. The science that uses solar oscillations to study the Sun is called **helioseismology**. Detailed observations of motions of material from place to place across the surface of the Sun show that the Sun vibrates or rings, something like a bell that has been struck. Unlike a well-tuned bell—which vibrates primarily at one frequency—the vibrations of the Sun are very complex. In the Sun, many different frequencies of vibrations occur simultaneously, which cause some parts of

Process of Science

LEARNING FROM FAILURE

The first detections of solar neutrinos raised more questions than they answered.



Newer laboratory and solar measurements confirmed the new hypothesis. Part of the "scientific attitude" is to find failure exciting. When experiments do not turn out as expected, good scientists get excited—there is something new to understand!

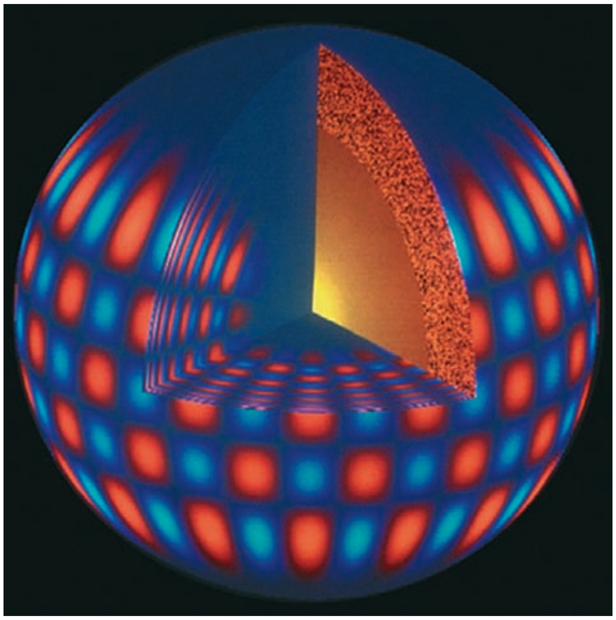


Figure 14.9 The interior of the Sun rings like a bell as helioseismic waves move through it. This figure shows one particular mode of the Sun's vibration. Red indicates regions where gas is traveling inward; blue indicates regions where gas is traveling outward. Astronomers observe these motions via Doppler shifts.

the Sun to bulge outward and some to draw inward. These motions help us to probe what lies below. **Figure 14.9** illustrates the motions of the different parts of the Sun, with red and blue areas moving in opposite directions. Some waves are amplified and some are suppressed, depending on how they overlap as they travel through the Sun. Astronomers study these waves using the Doppler effect (see Chapter 5), which distinguishes between parts of the Sun that move toward the observer and parts that move away.

To detect the disturbances of helioseismic waves on the surface of the Sun, astronomers must measure Doppler shifts of less than 0.1 m/s while detecting changes in brightness of only a few parts per million at any given location on the Sun. Tens of millions of different wave motions are possible within the Sun. Some waves travel around the circumference of the Sun, providing information about the density of the upper convection zone. Other waves travel through the interior of the Sun, revealing the density structure of the Sun close to its core. Still others travel inward toward the center of the Sun, until they are bent by the changing solar density and return to the surface.

All of these wave motions are going on at the same time, so sorting them out requires computer analysis of long, unbroken strings of solar observations from several sources. The Global Oscillation Network Group (GONG) is a network of six solar observation stations spread around the world that enables astronomers to observe the surface of the Sun approximately 90 percent of the time.

To interpret helioseismology data, scientists compare the measurements of the strength, frequency, and wavelengths of the waves against predicted vibrations calculated from models of the solar interior. This technique provides a powerful test of models of the solar interior, and it has led both to some surprises and to improvements in the models. For example, some scientists proposed that the solar neutrino problem might be solved if the models had overestimated the amount of helium in the Sun. This explanation was ruled out by analysis of the waves that penetrate to the core of the Sun. Helioseismology showed that the value for opacity used in early solar models was too low. This realization led astronomers to recalculate the location of the bottom of the convective zone. Both theory and observation now put the base of the convective zone at 70 percent of the way out from the center of the Sun, with an uncertainty in this number of less than half a percent.

Working back and forth between observation and theory has enabled astronomers to probe the otherwise inaccessible interior of the Sun. We now know that the energy is produced by nuclear fusion deep in the core and that it moves outward by radiation to a point about 70 percent of the radius of the Sun. Then it travels outward by convection to the surface. We also know how the temperature, density, and pressure change with radius and how these factors change the opacity at different distances from the center. Even though it is usually not possible to sample directly or to set up controlled experiments, this kind of collaboration between theory and observation is essential to observational sciences like astronomy.

CHECK YOUR UNDERSTANDING 14.3

How do neutrinos help us understand what is going on in the core of the Sun? (a) Neutrinos from distant objects pass through the Sun, probing the interior. (b) Neutrinos from the Sun pass easily through Earth. (c) Neutrinos from the interior of the Sun easily escape. (d) Neutrinos change form on their way to Earth.

14.3 The Atmosphere of the Sun

Beyond the convective zone lie the outer layers of the Sun, which are collectively known as the Sun's atmosphere. These layers, shown in **Figure 14.10a**, include the *photosphere*, the *chromosphere*, and the *corona*. We can observe these layers of the Sun directly using telescopes and satellites. Observations of the Sun's atmosphere are important because activity in the Sun's atmosphere has consequences for human infrastructure such as power grids and satellites in orbit around Earth.

The Sun is a large ball of gas, and so, unlike Earth, it has no solid surface. Its apparent surface is like a fog bank on Earth. Imagine watching some people walking into a fog bank. After they disappear from view, you would say they were inside the fog bank, even though they never passed through a definite boundary. The apparent surface of the Sun is similar. Light from the Sun's surface can escape directly into space, so we can see it. Light from below the Sun's surface cannot escape directly into space, so we cannot see it.

At the base of the atmosphere is the **photosphere**: the Sun's apparent surface. This is where features such as sunspots can be seen. Above this photosphere is the **chromosphere**, a region of strong emission lines. The top layer is the **corona**, which can be viewed during a solar eclipse as a halo around the Sun. In the Sun's atmosphere, the density of the gas drops very rapidly with increasing altitude. Figure 14.10b shows how density and temperature change across the atmosphere of the Sun. In this section, we will explore each of these layers, beginning at the bottom with the photosphere.

The Photosphere

The **effective temperature** of the photosphere is calculated from the Sun's luminosity and radius using the Stefan-Boltzmann law (see Chapter 5). The photosphere has an effective temperature of 5780 K, ranging from 6600 K to 4500 K over a 500-km-thick zone. As you can see in the graphs in Figure 14.10b, the temperature

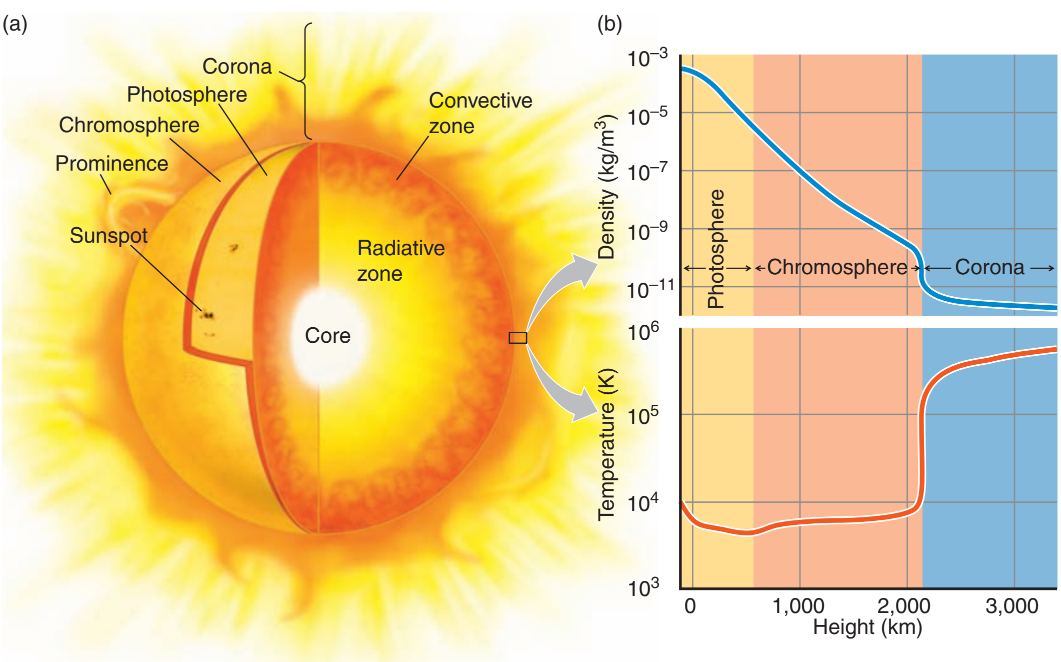


Figure 14.10 (a) The components of the Sun's atmosphere are located above the convective zone. (b) The density and temperature of the Sun's atmosphere change abruptly at the boundary between the chromosphere and corona. Note that the *y*-axes are logarithmic.

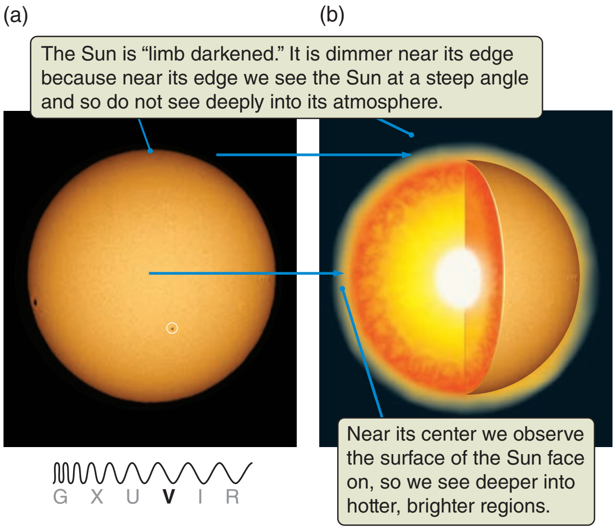


Figure 14.11 (a) When viewed in visible light, the Sun appears to have a sharp outline, even though it has no true surface. The center of the Sun appears brighter, while the limb of the Sun is darker—an effect known as limb darkening. (b) Looking at the center of the Sun allows us to see deeper into the Sun's interior than we do when looking at the edge of the Sun. Because higher temperature means more luminous radiation, the center of the Sun appears brighter than its limb.

increases sharply across the boundary between the chromosphere and the corona, while the density falls sharply across the same boundary. The Sun appears to have a well-defined surface and a sharp outline when viewed from Earth because 500 km does not look very thick when viewed from a distance of 150 million km.

In **Figure 14.11a**, the Sun appears fainter near its edges than near its center, an effect called **limb darkening**. This effect is an artifact of the structure of the Sun's photosphere. When looking near the edge of the Sun, you are looking through the photosphere at a steep angle. As a result, you do not see as deeply into the interior of the Sun as when you are looking directly down through the photosphere near the center of the Sun's disk. The light from the limb of the Sun comes from a shallower layer that is cooler and fainter, as shown in Figure 14.11b.

In the Sun's atmosphere, the density of the gas drops very rapidly with increasing altitude. All visible solar phenomena take place in the Sun's atmosphere. Most of the radiation from below the Sun's photosphere is absorbed by matter and reemitted at the photosphere as a blackbody spectrum.

As we examine the structure of the Sun in more detail, however, we see that this simple description of the spectra of stars is incomplete. Light from the solar photosphere must escape through the upper layers of the Sun's atmosphere, which affects the spectrum we observe. In Chapter 13, we discussed the presence of absorption lines in the spectra of stars. Now we can take a closer look at how these absorption lines form. As photospheric light travels upward, atoms in the solar atmosphere absorb the light at discrete wavelengths, forming absorption lines. Because the Sun appears so much brighter than any other star, its spectrum can be studied in far more detail, so specially designed telescopes and high-resolution spectrometers have been built specifically to study the Sun's light. The solar spectrum is shown in **Figure 14.12**. Absorption lines from more than 70 elements have been identified. Analysis of these lines forms the basis for much of astronomers' knowledge of the solar atmosphere, including the composition of the Sun. This is also the starting point for an understanding of the atmospheres and spectra of other stars.

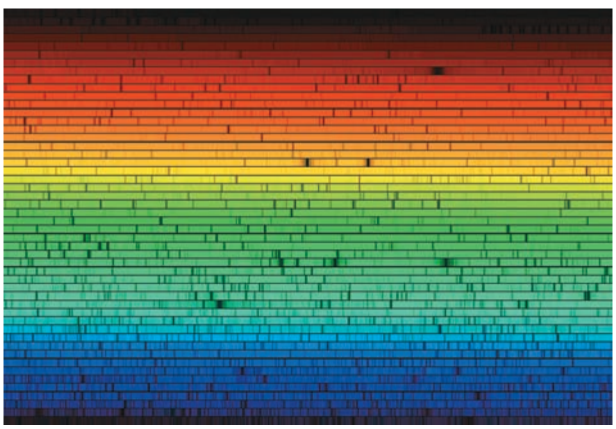


Figure 14.12 This high-resolution spectrum of the Sun stretches from 400 nanometers (nm) in the lower left corner to 700 nm in the upper right corner and shows dark absorption lines. This spectrum was produced by passing the Sun's light through a prism-like device, and then cutting and folding the single long spectrum (from blue to red) into rows so that it will fit in a single image taken by a camera.

The Chromosphere and Corona

Moving upward through the Sun's photosphere, the temperature falls from 6600 K at the photosphere's bottom to 4400 K at its top. At this point, the trend reverses and the temperature slowly begins to climb, rising to about 6000 K at a height of 1,500 km above the top of the photosphere (see Figure 14.10b). This region of increasing temperature is called the chromosphere (**Figure 14.13a**). The reason for the chromosphere's temperature reversal with increasing height is not well understood, but it may be caused by magnetic waves propagating through the region and depositing their energy at the top of the chromosphere.

The chromosphere was discovered in the 19th century during observations of total solar eclipses (Figure 14.13b). The chromosphere is seen most strongly at the solar limb as a source of emission lines, especially a particular hydrogen line that is produced when the electron falls from the third energy state to the second energy state. This line is known as the Hα line (the "hydrogen alpha line"). The deep red color of the Hα line is what gives the chromosphere its name; the word means "the place where color comes from." The element helium was discovered in 1868 from a spectrum of the chromosphere of the Sun nearly 30 years before it was found on Earth: helium is named after *helios*, the Greek word for "Sun."

At the top of the chromosphere, across a transition region that is only about 100 km thick, the temperature suddenly soars while the density abruptly drops

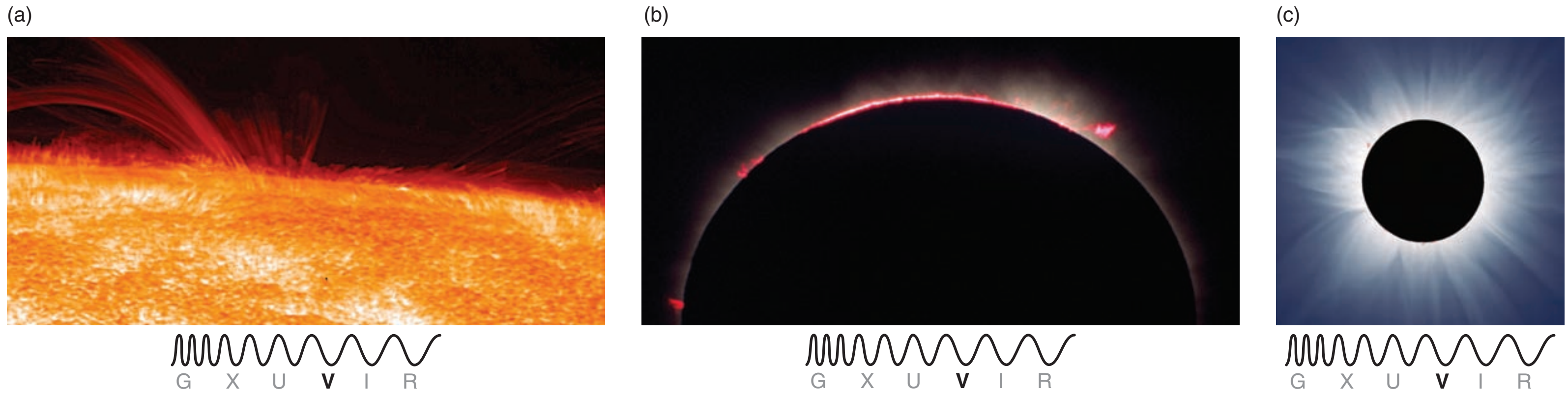


Figure 14.13 (a) This spacecraft image of the Sun shows fine structure in the chromosphere extending outward from the photosphere. (b) The chromosphere is visible during a total eclipse. (c) This eclipse image shows the Sun's corona, consisting of million-kelvin gas that extends for millions of kilometers beyond the surface of the Sun.

(see Figure 14.10b). Above this transition lies the outermost region of the Sun's atmosphere, the corona, where temperatures reach 1 million to 2 million K. The corona is thought to be heated by magnetic fields and micro solar flares.

The Sun's corona has been known since ancient times: it is visible during total solar eclipses as an eerie outer glow stretching a distance of several solar radii beyond the Sun's surface (Figure 14.13c). Because it is so hot, the solar corona is also a strong source of X-rays, and there is so much energy in these X-ray photons that many electrons are stripped away from nuclei, leaving atoms in the corona highly ionized.

CHECK YOUR UNDERSTANDING 14.4

The surface of the Sun appears sharp in visible light because: (a) the photosphere is cooler than the layers below it; (b) the photosphere is thin compared to the other layers in the Sun; (c) the photosphere is less dense than the convection zone; (d) the Sun has a distinct surface.

14.4 The Atmosphere of the Sun Is Very Active

The atmosphere of the Sun is a very turbulent place. The best-known features on the surface of the Sun are relatively dark blemishes in the solar photosphere, called **sunspots**. Sunspots come and go over time, though they remain long enough for us to determine the rotation rate of the Sun. These spots are associated with active regions: loops of material and explosions that fling particles far out into the Solar System. Long-term patterns have been observed in the variations of sunspots and active regions, revealing that the magnetic field of the Sun is constantly changing.

Solar Activity Is Caused by Magnetic Effects

The magnetic field (see Chapter 5) of the Sun causes virtually all of the structure seen in the Sun's atmosphere. High-resolution images of the Sun show *coronal loops* that make up much of the Sun's lower corona (**Figure 14.14**). This texture is the result of magnetic structures called flux tubes. Magnetic fields are responsible for much of the structure of the corona as well. The corona is far too hot to be held in by the Sun's gravity, but over most of the surface of the Sun, coronal gas is

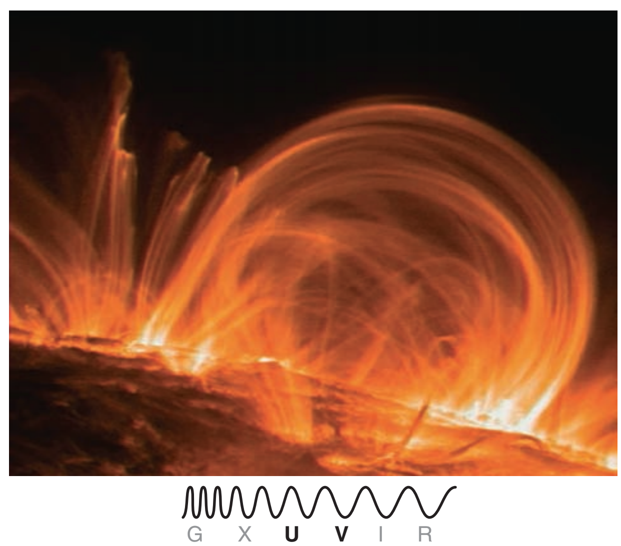


Figure 14.14 This close-up image of the Sun shows the tangled structure of coronal loops.

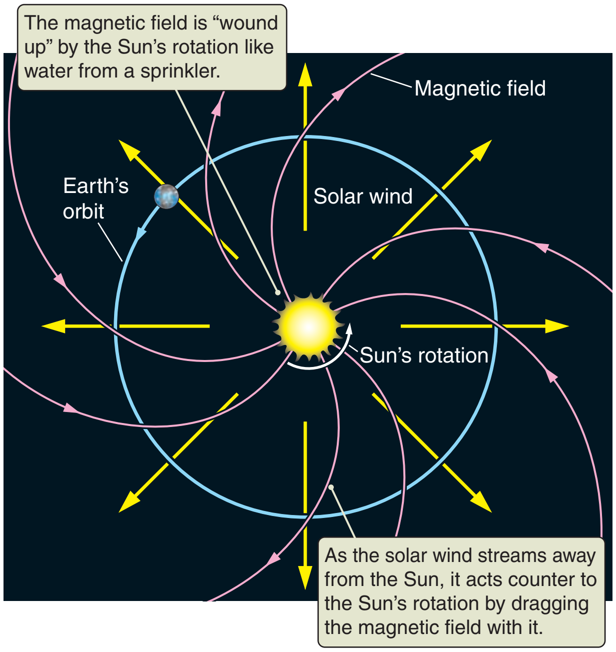


Figure 14.15 The solar wind streams away from active areas and coronal holes on the Sun. As the Sun rotates, the solar wind takes on a spiral structure, much like the spiral of water that streams away from a rotating lawn sprinkler.

confined by magnetic loops with both ends firmly anchored deep within the Sun. The magnetic field in the corona acts almost like a network of rubber bands that coronal gas is free to slide along but cannot cross. In contrast, about 20 percent of the surface of the Sun is covered by an ever-shifting pattern of **coronal holes**, which are large regions where the magnetic field points outward, away from the Sun, and where coronal material is free to stream away into interplanetary space as the solar wind. In extreme ultraviolet images of the Sun we see coronal holes as dark regions, which indicates that they are cooler and lower in density than their surroundings (see the chapter-opening photograph).

The relatively steady part of the solar wind consists of lower-speed flows with velocities of about 350 km/s and higher-speed flows with velocities up to about 700 km/s. The higher-speed flows originate in coronal holes. Depending on their speed, particles in the solar wind take about 2–5 days to reach Earth. Frequently, 2–5 days after a coronal hole passes across the center of the face of the Sun, the speed and density of the solar wind reaching Earth increases. As you can see in **Figure 14.15**, the solar wind drags the Sun's magnetic field along with it. The magnetic field in the solar wind gets "wound up" by the Sun's rotation. Consequently, the solar wind has a spiral structure resembling the stream of water from a rotating lawn sprinkler.

The effects of the solar wind are felt throughout the Solar System. The solar wind blows the tails of comets away from the Sun, shapes the magnetospheres of the planets, and provides the energetic particles that power Earth's spectacular auroral displays. Using space probes, astronomers have been able to observe the solar wind extending out to 100 astronomical units (AU) from the Sun. But the solar wind does not go on forever. The farther it gets from the Sun, the more it has to spread out. Just like radiation, the density of the solar wind follows an inverse square law. At a distance of about 100 AU from the Sun, the solar wind stops abruptly. Here it piles up against the pressure of the **interstellar medium**, which is the gas and dust that lie between stars in a galaxy. **Figure 14.16** shows the

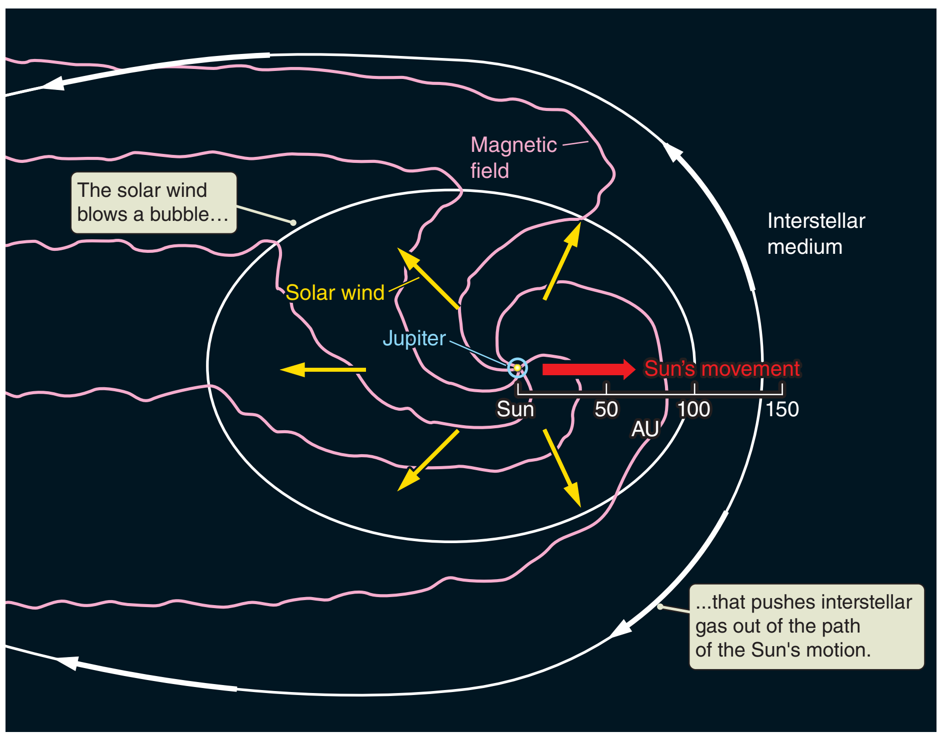


Figure 14.16 The solar wind streams away from the Sun for about 100 AU, until it finally piles up against the pressure of the interstellar medium through which the Sun is traveling. The *Voyager 1* spacecraft has recently crossed this boundary.

region of space over which the solar wind is measured. The *Voyager 1* spacecraft has crossed the outer edge of this boundary and sent back the first direct measurements of true interstellar space. The *Interstellar Boundary Explorer* spacecraft, launched in 2008, is also exploring this region.

Sunspots and Changes in the Sun

Sunspots have been noted since antiquity. Telescopic observations of sunspots date back almost 400 years, and there are records of naked-eye observations by Chinese, Greek, and medieval astronomers centuries before that. *But remember that you should never look directly at the Sun!* Direct viewing through a commercial solar filter is safe, as is projecting the image through a telescope or binoculars onto a surface such as paper and looking only at the projection. Many websites have live images of the Sun viewed through ground and space telescopes (see the "Using the Web" problems at the end of the chapter). Sunspots are places where material is trapped at the surface of the Sun by magnetic-field lines. When this material cools, convection cannot carry it downward, so it makes a cooler (therefore darker) spot on the surface of the Sun. **Figure 14.17** shows a large sunspot group. Sunspots appear dark, but only in contrast to the brighter surface of the Sun **(Working It Out 14.2)**.

Early telescopic observations of sunspots made during the 17th century led to the discovery of the Sun's rotation, which has an average period of about 27 days as seen from Earth and 25 days relative to the stars. Because Earth orbits the Sun in the same direction that the Sun rotates, observers on Earth see a slightly longer rotation period. Observations of sunspots also show that the Sun's rotation period

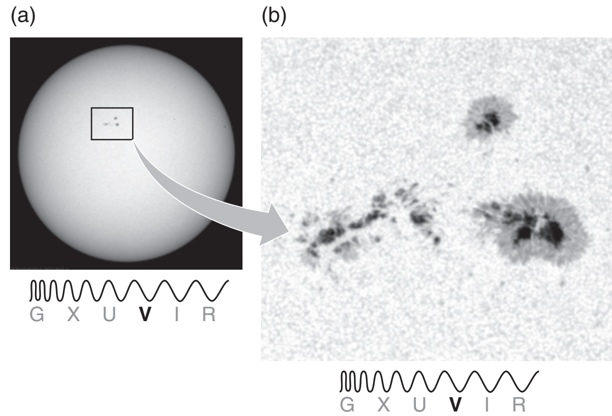


Figure 14.17 (a) This image from the Solar Dynamics Observatory (SDO), taken in 2010, shows a large sunspot group. Sunspots are magnetically active regions that are cooler than the surrounding surface of the Sun. (b) This high-resolution view shows the sunspots in this group.

14.2 Working It Out Sunspots and Temperature

Sunspots are about 1500 K cooler than their surroundings. What does this lower temperature tell us about their luminosity? Think back to the Stefan-Boltzmann law in Chapter 5. The flux, $\mathcal{F}$, from a blackbody is proportional to the fourth power of the temperature, T. The constant of proportionality is the Stefan-Boltzmann constant, σ, which has a value of 5.67×10^{-8} W/(m^2 K^4). We write this relationship as

$$\mathcal{F} = \sigma T^4$$

Remember that the flux is the amount of energy coming from a square meter of surface every second. How much less energy comes out of a sunspot than out of the rest of the Sun? Let's take round numbers for the temperature of a typical sunspot and the surrounding photosphere: 4500 K and 6000 K, respectively. We can set up two equations:

$$\mathcal{F}_{\text{spot}} = \sigma T^4_{\text{spot}} \quad \text{and} \quad \mathcal{F}_{\text{surface}} = \sigma T^4_{\text{surface}}$$

We could solve each of these separately, and then divide the value of $\mathcal{F}_{\text{spot}}$ by $\mathcal{F}_{\text{surface}}$ to find out how much fainter the sunspot is, but it's much easier to solve for the ratio of the fluxes:

$$\frac{\mathcal{F}_{\text{spot}}}{\mathcal{F}_{\text{surface}}} = \frac{\sigma T^4_{\text{spot}}}{\sigma T^4_{\text{surface}}} = \frac{T^4_{\text{spot}}}{T^4_{\text{surface}}} = \left(\frac{T_{\text{spot}}}{T_{\text{surface}}}\right)^4$$

Plugging in our values for T_{spot} and T_{surface} gives

$$\frac{\mathcal{F}_{\text{spot}}}{\mathcal{F}_{\text{surface}}} = \left(\frac{4500\text{ K}}{6000\text{ K}}\right)^4 = 0.32$$

and multiplying both sides by $\mathcal{F}_{\text{surface}}$ gives

$$\mathcal{F}_{\text{spot}} = 0.32\mathcal{F}_{\text{surface}}$$

So the amount of energy coming from a square meter of sunspot every second is about one-third as much as the amount of energy coming from a square meter of surrounding surface every second. In other words, the sunspot is about one-third as bright as the surrounding photosphere. If you could cut out the sunspot and place it elsewhere in the sky, it would be brighter than the full Moon.

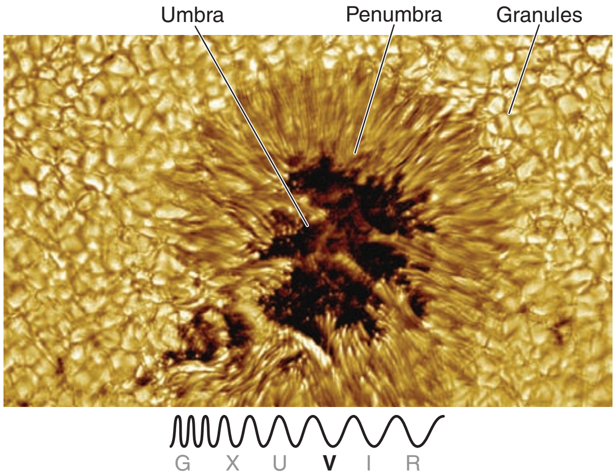


Figure 14.18 This very high-resolution view of a sunspot shows the dark umbra surrounded by the lighter penumbra. The solar surface around the sunspot bubbles with separate cells of hot gas called *granules*. The smallest features are about 100 km across.

is shorter at its equator than at higher latitudes, an effect called **differential rotation**. Differential rotation is possible only because the Sun is a large ball of gas rather than a solid object.

Figure 14.18 shows the structure of a sunspot on the surface of the Sun. A sunspot consists of an inner dark core called the **umbra**, which is surrounded by a less dark region called the **penumbra**, which shows an intricate radial pattern, reminiscent of the petals of a flower. Sunspots are caused by magnetic fields thousands of times greater than the magnetic field at Earth's surface. They occur in pairs that are connected by loops in the magnetic field. Sunspots range in size from a few tens of kilometers across up to complex groups that may contain several dozen individual spots and span as much as 150,000 km. The largest sunspot groups are so large that they can be seen without a telescope.

Although sunspots occasionally last 100 days or longer, half of all sunspots come and go in about 2 days, and 90 percent are gone within 11 days. The number and distribution of sunspots change over time in a pattern averaging 11 years called the **sunspot cycle**. **Figure 14.19a** shows data for several recent cycles. At the beginning of a cycle, sunspots appear at solar latitudes of about 30° north and south of the solar equator. Over the following years, sunspots are found closer to the equator as their number increases to a maximum and then declines.

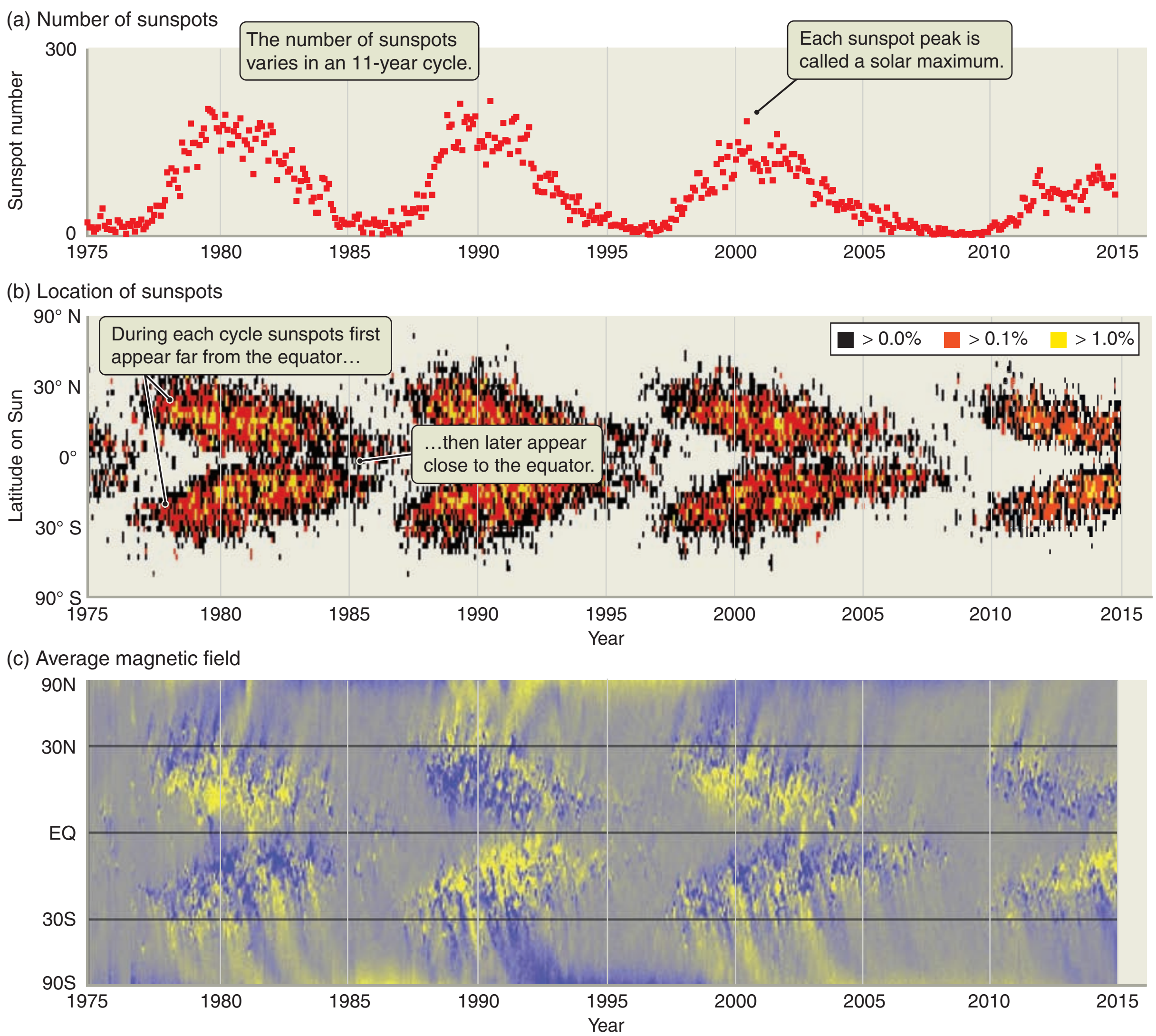


Figure 14.19 (a) The number of sunspots varies with time, as shown in this graph of the past few solar cycles. (b) The "solar butterfly" diagram shows the fraction of the Sun covered by sunspots at each latitude. The data are color coded to show the percentage of the strip at that latitude that is covered in sunspots at that time: black, 0 to 0.1 percent; red, 0.1–1.0 percent; yellow, greater than 1.0 percent. (c) The Sun's magnetic poles flip every 11 years. Yellow indicates magnetic south; blue indicates magnetic north.

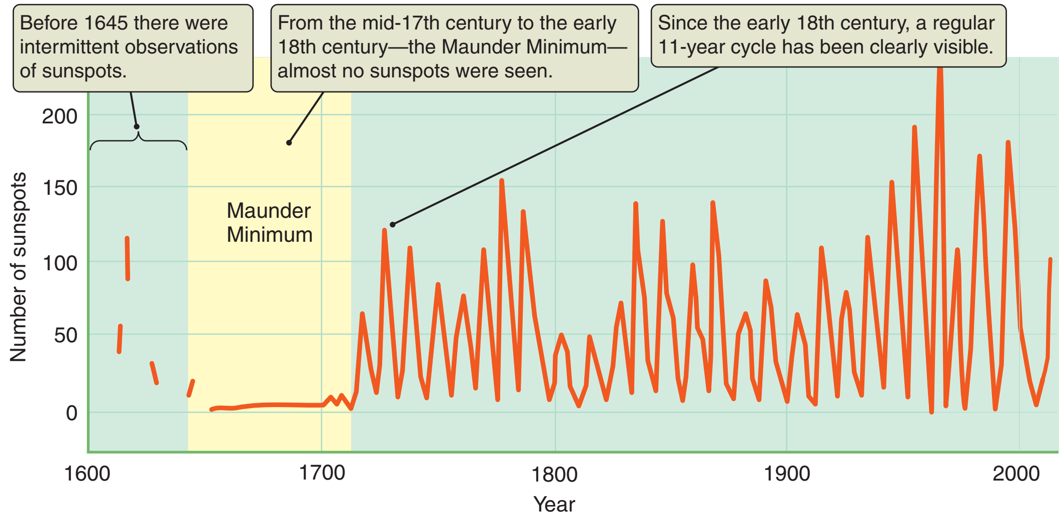


Figure 14.20 Sunspots have been observed for hundreds of years. In this plot, the 11-year cycle in the number of sunspots (half of the 22-year solar magnetic cycle) is clearly visible. Sunspot activity varies greatly over time. The period from the middle of the 17th century to the early 18th century, when almost no sunspots were seen, is called the *Maunder Minimum*.

As the last few sunspots approach the equator, new sunspots again begin appearing at middle latitudes, and the next cycle begins. Figure 14.19b shows the number of sunspots at a given latitude plotted against time: this diagram of opposing diagonal bands is called the sunspot "butterfly diagram."

In the early 20th century, solar astronomer George Ellery Hale (1868–1938) was the first to show that the 11-year sunspot cycle is actually half of a 22-year magnetic cycle during which the direction of the Sun's magnetic field reverses after each 11-year sunspot cycle. Figure 14.19c shows how the average strength of the magnetic field at every latitude has changed over more than 35 years. The direction of the Sun's magnetic field flips at the maximum of each sunspot cycle. Sunspots tend to come in pairs, with one spot (the leading sunspot) in front of the other with respect to the Sun's rotation. In one 11-year sunspot cycle, the leading sunspot in each pair tends to be a north magnetic pole, whereas the trailing sunspot tends to be a south magnetic pole. In the next 11-year sunspot cycle, this polarity is reversed: the leading sunspot in each pair is a south magnetic pole, whereas the trailing sunspot tends to be a north magnetic pole. The transition between these two magnetic polarities occurs near the peak of each sunspot cycle. Magnetic activity on the Sun affects the photosphere, chromosphere, and corona.

Telescopic observations of sunspots date back almost 400 years. As you can see in **Figure 14.20**, the 11-year cycle is neither perfectly periodic nor especially reliable. The time between peaks in the number of sunspots actually varies between about 9.7 and 11.8 years. The number of spots seen during a given cycle fluctuates as well, and there have been periods when sunspot activity has disappeared almost entirely. An extended lull in solar activity, called the **Maunder Minimum**, lasted from 1645 to 1715. Typically, there are about six peaks of solar activity in 70 years, but virtually no sunspots were seen during the Maunder Minimum, and auroral displays were less frequent than usual.

Sunspots are only one of several phenomena that follow the Sun's 22-year cycle of magnetic activity. The peaks of the cycle, called **solar maxima**, are times of intense activity. Sunspots are often accompanied by a brightening of the solar chromosphere that is seen most clearly in emission lines such as Hα. These bright regions are known as solar active regions. The magnificent loops arching through the solar corona, shown in **Figure 14.21**, are solar **prominences**, magnetic flux tubes of relatively cool (5000–10,000 K) but dense gas extending through the

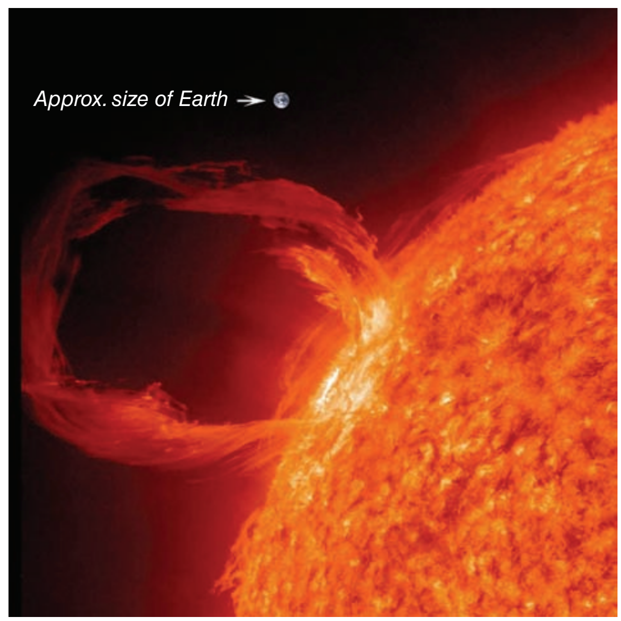


Figure 14.21 Solar prominences are magnetically supported arches of hot gas that rise high above active regions on the Sun. Here, you can see a close-up view at the base of a large prominence. An image of Earth is included for scale (it is not actually that close to the Sun).

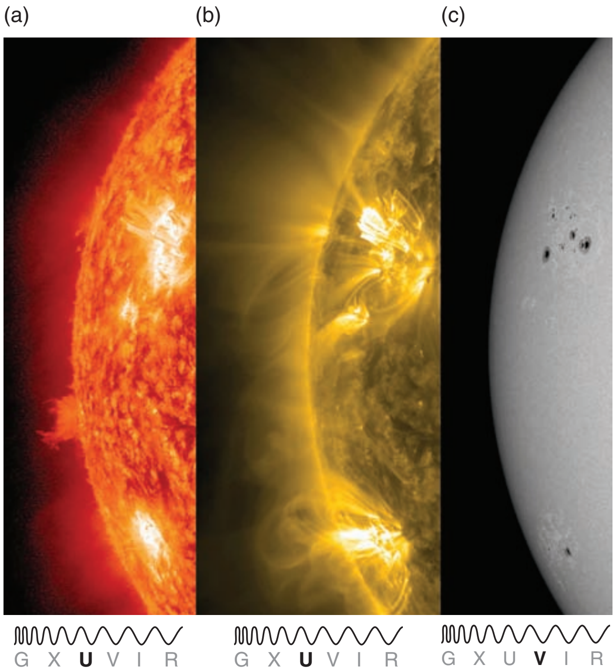


Figure 14.22 The Solar Dynamics Observatory (SDO) observed these active regions of the Sun that produced solar flares in August 2011. (a) Activity near the surface at 60,000 K is visible in extreme ultraviolet light (along with a prominence rising up from the Sun's edge). (b) Viewed at other ultraviolet wavelengths, many looping arcs and plasma heated to about 1 million K become visible. (c) The dark spots in this image are the magnetically intense sunspots that are the sources of all the activity.

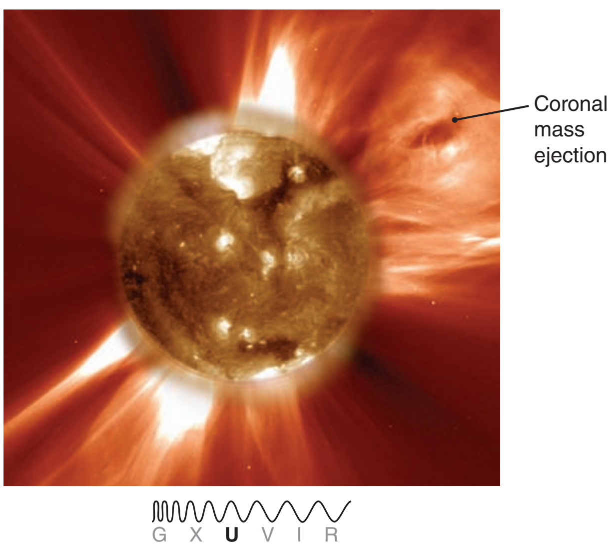


Figure 14.23 This Solar and Heliospheric Observatory (*SOHO*) image shows a coronal mass ejection (upper right): a simultaneously recorded ultraviolet image of the solar disk is superimposed.

million-kelvin gas of the corona. These prominences are anchored in the active regions. Although most prominences are relatively quiet, others can erupt out through the corona, towering a million kilometers or more over the surface of the Sun and ejecting material into the corona at velocities of 1,000 km/s.

Solar flares are the most energetic form of solar activity, violent eruptions in which enormous amounts of magnetic energy are released over the course of a few minutes to a few hours. **Figure 14.22** shows solar flares erupting from two sunspot groups. The two left-hand images (Figures 14.22a and b), taken in ultraviolet light, show material at very high temperatures. The spots in the visible-light image on the right (Figure 14.22c) are at the base of the activity seen in Figures 14.22a and b. Solar flares can heat gas to temperatures of 20 million K, and they are the source of intense X-rays and gamma rays. Hot **plasma** (consisting of atoms stripped of some of their electrons) moves outward from flares at speeds that can reach 1,500 km/s. Magnetic effects can then accelerate subatomic particles to almost the speed of light. Such events, called **coronal mass ejections (CMEs)** (**Figure 14.23**), send powerful bursts of energetic particles outward through the Solar System. Coronal mass ejections occur about once per week during the minimum of the sunspot cycle and as often as several times per day near the maximum of the cycle.

Solar Activity Affects Earth

The amount of solar radiation received at the distance of Earth from the Sun has been measured to be, on average, 1,361 watts per square meter (W/m^2). As you can see in **Figure 14.24**, satellite measurements of the amount of radiation coming from the Sun show that this value varies by as much as 0.2 percent over periods of a few weeks, as dark sunspots in the photosphere and bright spots in the chromosphere move across the disk. Overall, however, the increased radiation from active regions on the Sun more than makes up for the reduction in radiation from sunspots. On average, the Sun seems to be about 0.1 percent brighter during the peak of a solar cycle than it is at its minimum.

Solar activity affects Earth in many ways. Solar active regions are the source of most of the Sun's extreme ultraviolet and X-ray emissions, energetic radiation that heats Earth's upper atmosphere and, during periods of increased solar activity, causes Earth's upper atmosphere to expand. When this happens, the swollen upper atmosphere can significantly increase the atmospheric drag on spacecraft orbiting at relatively low altitudes, such as that of the Hubble Space Telescope, causing their orbits to decay. Periodic boosts have been necessary to keep the Hubble Space Telescope in its orbit.

Earth's magnetosphere is the result of the interaction between Earth's magnetic field and the solar wind. Increases in the solar wind accompanying solar activity, especially coronal mass ejections directed at Earth, can disrupt Earth's magnetosphere. Spectacular auroras can accompany such events, as can magnetic storms that have been known to disrupt electric power grids and cause blackouts across large regions. Coronal mass ejections that are emitted in the direction of Earth also hinder radio communication and navigation, and they can damage sensitive satellite electronics, including communication satellites. In addition, energetic particles accelerated in solar flares pose one of the greatest dangers to human exploration of space.

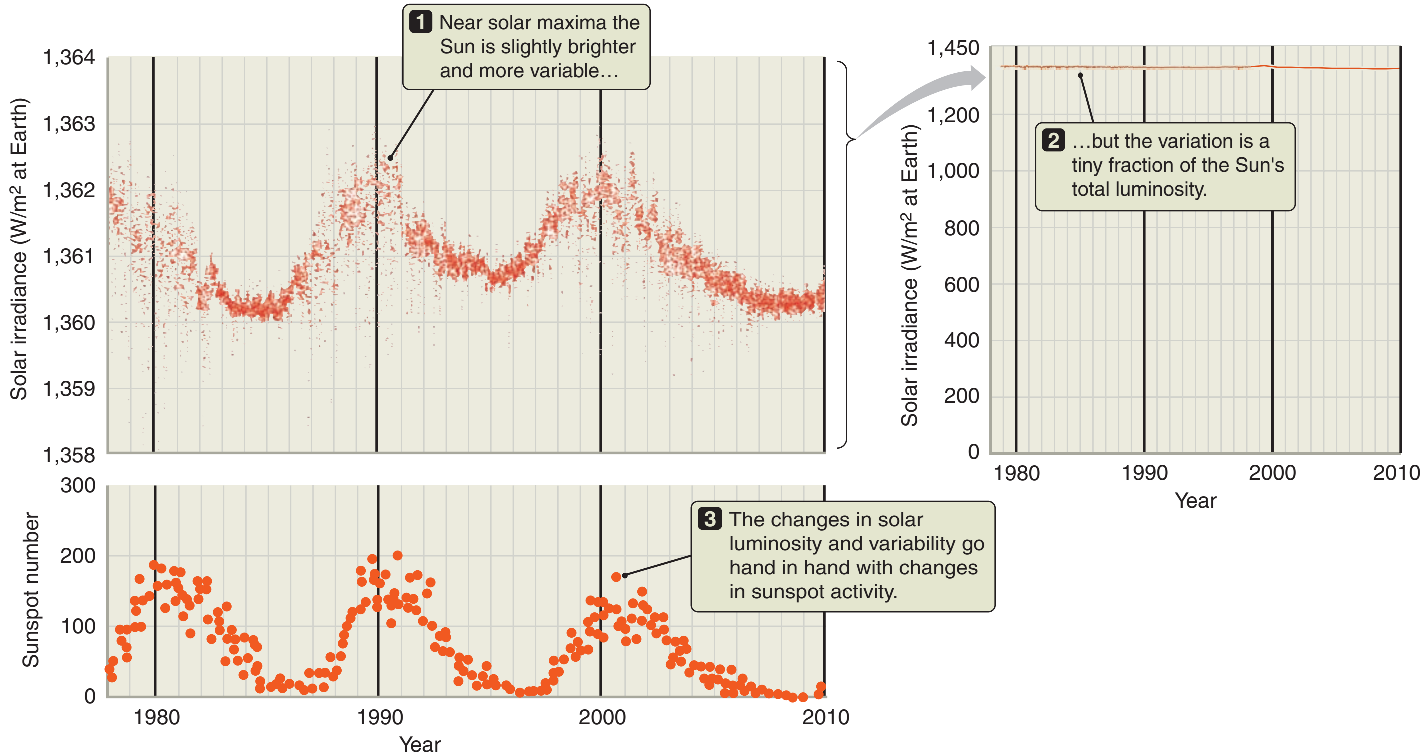


Figure 14.24 Measurements taken by satellites above Earth's atmosphere show that the amount of light from the Sun changes slightly over time.

Detailed observations from the ground and from space help astronomers understand the complex nature of the solar atmosphere. The Solar and Heliospheric Observatory (*SOHO*) spacecraft is a joint mission between NASA and the European Space Agency (ESA). *SOHO* moves in lockstep with Earth at a location approximately 1,500,000 km from Earth that is almost directly in line between Earth and the Sun. *SOHO* carries 12 scientific instruments that monitor the Sun and measure the solar wind upstream of Earth. Additionally, NASA's Solar Dynamics Observatory (SDO) studies the solar magnetic field in order to predict when major solar events will occur, rather than simply responding after they happen.

CHECK YOUR UNDERSTANDING 14.5

Sunspots appear dark because: (a) they have very low density; (b) magnetic fields absorb most of the light that falls on them; (c) they are regions of very high pressure; (d) they are cooler than their surroundings.

Origins

The Solar Wind and Life

Solar flares and coronal mass ejections can affect the space around Earth. In fact, energetic particles accelerated in solar flares pose one of the greatest dangers to human exploration of space, and they need to be considered when astronauts are orbiting Earth in a space station or, someday, traveling to the Moon or farther. Earth's magnetic field protects life on the surface from these energetic particles: the particles travel along the magnetic-field lines to Earth: poles, creating the auroras, without bombarding the surface, which could be harmful to life on Earth. But the Moon does not have this protection because its magnetic field is very weak. Astronauts on the lunar surface would be exposed to as much radiation as astronauts traveling in space. The strength of the solar wind varies with the solar cycle, as noted in Section 14.4, so exposure danger varies as well.

As illustrated in **Figure 14.25**, the Solar System is surrounded by the **heliosphere**, in which the solar wind blows against the interstellar medium and clears out an area like the inside of a bubble. As the Sun and Solar System move through the Milky Way Galaxy, passing in and out of interstellar clouds, this heliosphere protects the entire Solar System from galactic high-energy particles known as cosmic rays that originate primarily in high-energy explosions of massive dying stars. When the Sun is in its lower-activity state, the heliosphere is weaker, so more galactic cosmic rays enter the Solar System. In addition, the intensity of these cosmic rays depends on where the Sun and Solar System are located in their orbit about the center of the Milky Way Galaxy.

Some scientists have theorized that at times when the Sun was quiet and the heliosphere was weaker than average, and the Solar System was passing through a particular part of the galaxy, the cosmic-ray flux in the Solar System—and on Earth—increased. This increased flux possibly led to a disruption in Earth's ozone layer and possibly contributed to a mass extinction in which many species died out on Earth.

Thus, in addition to the obvious contribution of the Sun to heat and light on Earth, the extension of the Sun through the solar wind may have affected the evolution of life on Earth—and it may also affect the ability of humans to live and work in space.

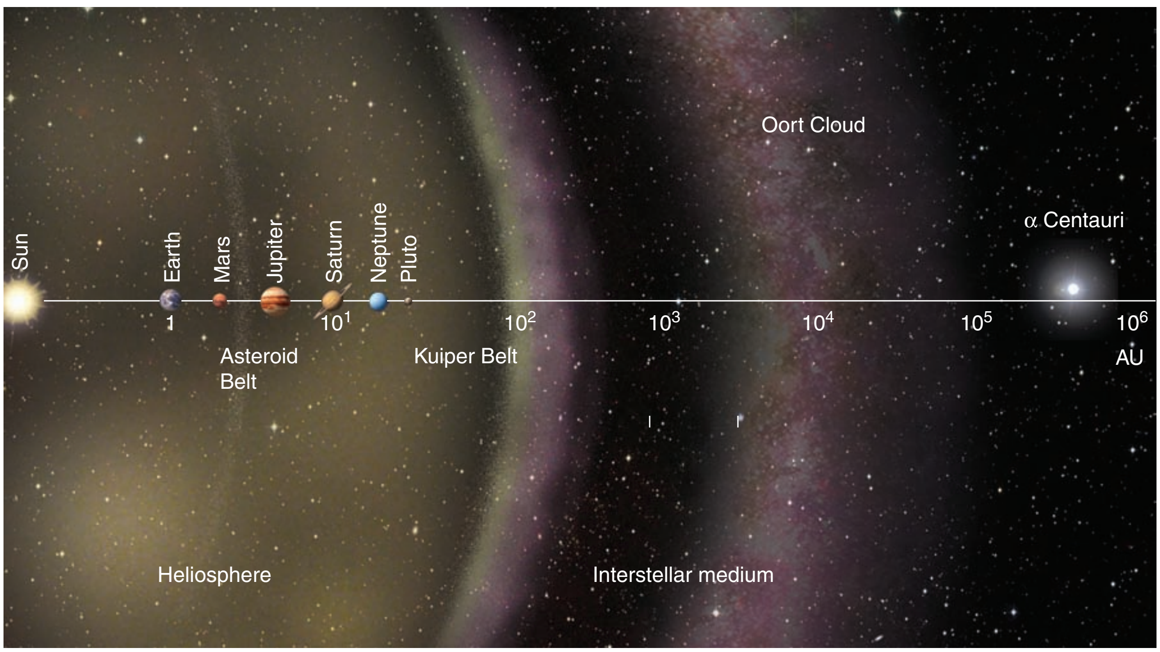


Figure 14.25 The heliosphere of the Sun, a bubble of charged particles partially covering the Solar System that is formed by the solar wind blowing against the interstellar medium. The *Voyager* spacecraft are just past 100 AU. Notice that the scale is logarithmic.

ARTICLES QUESTIONS

In this article, scientists discuss major solar storms that affect Earth.

Carrington-Class CME Narrowly Misses Earth

By **DR. TONY PHILLIPS, Science@NASA**

Last month (April 8–11, 2014), scientists, government officials, emergency planners, and others converged on Boulder, Colorado, for NOAA's Space Weather Workshop—an annual gathering to discuss the perils and probabilities of solar storms.

The current solar cycle is weaker than usual, so you might expect a correspondingly low-key meeting. On the contrary, the halls and meeting rooms were abuzz with excitement about an intense solar storm that narrowly missed Earth.

"If it had hit, we would still be picking up the pieces," says Daniel Baker of the University of Colorado, who presented a talk entitled The Major Solar Eruptive Event in July 2012: Defining Extreme Space Weather Scenarios.

The close shave happened almost two years ago. On July 23, 2012, a plasma cloud or "CME" rocketed away from the Sun as fast as 3,000 km/s, more than four times faster than a typical eruption. The storm tore through the Earth's orbit, but fortunately Earth wasn't there. Instead it hit the *STEREO-A* spacecraft. Researchers have been analyzing the data ever since, and they have concluded that the storm was one of the strongest in recorded history. "It might have been stronger than the Carrington Event itself," says Baker.

The Carrington Event of September 1859 was a series of powerful CMEs that hit Earth head-on, sparking Northern Lights as far south as Tahiti. Intense geomagnetic storms caused global telegraph lines to spark, setting fire to some telegraph offices and disabling the "Victorian Internet." A similar storm today could have a catastrophic effect on modern power grids and telecommunication networks. According to a study by the National Academy of Sciences, the total economic impact could exceed $2 trillion or 20 times greater than the costs of a Hurricane Katrina. Multi-ton transformers fried by such a storm could take years to repair and impact national security.

A recent paper in *Nature Communications* authored by UC Berkeley space physicist Janet G. Luhmann and former postdoc Ying D. Liu describes what gave the July 2012 storm Carrington-like potency. For one thing, the CME was actually *two* CMEs separated by only 10 to 15 minutes. This double storm cloud traveled through a region of space that had been cleared out by another CME only four days earlier. As a result, the CMEs were not decelerated as much as usual by their transit through the interplanetary medium.

Had the eruption occurred just one week earlier, the blast site would have been facing Earth, rather than off to the side, so it was a relatively narrow escape.

When the Carrington Event enveloped Earth in the 19th century, technologies of the day were hardly sensitive to electromagnetic disturbances. Modern society, on the other hand, is deeply dependent on Sun-sensitive technologies such as GPS, satellite communications, and the Internet.

"The effect of such a storm on our modern technologies would be tremendous," says Luhmann.

During informal discussions at the workshop, Nat Gopalswamy of the Goddard Space Flight Center noted that "without NASA's *STEREO* probes, we might never have known the severity of the 2012 superstorm. This shows the value of having 'space weather buoys' located all around the Sun."

It also highlights the potency of the Sun even during so-called "quiet times." Many observers have noted that the current solar cycle is weak, perhaps the weakest in 100 years. Clearly, even a weak solar cycle can produce a very strong storm. Says Baker, "We need to be prepared."

ARTICLES QUESTIONS

1. What is a CME?
2. Why would a CME cause disruptions on Earth?
3. Explain how this storm missed Earth by 1 week.
4. More sensationalistic headlines for this story claimed that Earth almost "was sent back to the Dark Ages." What did they mean by this exaggeration?
5. Go to the NASA press release for this story (http://science.nasa.gov/science-news/science-at-nasa/2014/23jul_superstorm/), and watch the 4-minute "ScienceCast" video. What happened during the Carrington CME in 1859? Is this video effective at communicating the science information to the nonspecialist?

Summary

The forces due to pressure and gravity balance each other in hydrostatic equilibrium, maintaining the Sun's structure. Nuclear reactions converting hydrogen to helium are the source of the Sun's energy. Energy created in the Sun's core moves outward to the surface, first by radiation and then by convection. The solar wind may adversely affect astronauts located in space or on planets that lack a protective magnetic field, but it also has protected the Solar System from galactic, high-energy, cosmic-ray particles.

LG 1 **Describe the balance between the forces that determine the structure of the Sun.** The outward pressure of the hot gas inside the Sun balances the inward pull of gravity at every point. This balance is dynamically maintained. An energy balance is also maintained, with the energy produced in the core of the Sun balancing the energy lost from the surface.

LG 2 **Explain how mass is converted into energy in the Sun's core and how long it will take the Sun to use up its fuel.** In the core of the Sun, mass is converted to energy via the proton-proton cycle. When four hydrogen atoms fix to one helium atom, some mass is lost. This mass is released as energy, nearly all of which leaves the Sun either as photons or as neutrinos. Neutrinos are elusive, almost massless particles that interact only very weakly with other matter. Observations of neutrinos confirm that nuclear fusion is the Sun's primary energy source.

LG 3 **Sketch a physical model of the Sun's interior, and list the different ways that energy moves outward from the Sun's core toward its surface.** The interior of the Sun is divided into zones that are defined by how energy is transported in that region. Energy moves outward through the Sun by radiation and by convection.

LG 4 **Describe how observations of solar neutrinos and seismic vibrations on the surface of the Sun test astronomers' models of the Sun.** The Sun has multiple layers, each with a characteristic pressure, density, and temperature. Neutrinos directly probe the interior of the Sun. This model of the interior of the Sun has been tested by helioseismology, in much the same way that the model of Earth's interior has been tested by seismology.

LG 5 **Describe the solar activity cycles of 11 and 22 years, and explain how these cycles are related to the Sun's changing magnetic field.** Activity on the Sun follows a cycle that peaks every 11 years but takes 22 full years for the magnetic field to reverse. Sunspots are photospheric regions that are cooler than their surroundings, and they reveal the cycles in solar activity. Material streaming away from the Sun's corona creates the solar wind, which moves outward through the Solar System until it meets the interstellar medium. Solar storms, including ejections of mass from the corona, produce auroras and can disrupt power grids and damage satellites.

? UNANSWERED QUESTIONS

- Will nuclear fusion become a major source of energy production on Earth? Scientists have been working on controlled nuclear fusion for more than 60 years, since the first hydrogen bombs were developed. But so far there are too many difficulties in replicating the conditions inside the Sun. Nuclear fusion requires that we have hydrogen isotopes at very high temperature, density, and pressure, just as is the case when a hydrogen bomb explodes. However, controlled nuclear fusion requires that we confine this material sufficiently long to get more energy out than we put in. Several major experiments have attempted to fuse isotopes of hydrogen. An alternative approach is to fuse an isotope of helium, ^{3}He, which has only three particles in the nucleus (two protons and one neutron). On Earth, ^{3}He is found in very limited supply. But ^{3}He is in much greater abundance on the Moon, so some people propose setting up mining colonies on the Moon to extract ^{3}He for use in fusion reactions on Earth or possibly even on the Moon (see question 49b at the end of the chapter).
- Are variations in Earth's climate related to solar activity? Solar activity affects Earth's upper atmosphere, and it may affect weather patterns as well. It has also been suggested that variations in the amount of radiation from the Sun might be responsible for past variations in Earth's climate. Current models indicate that observed variations in the Sun's luminosity could account for only about 0.1 K differences in Earth's average temperature—much less than the effects due to the ongoing buildup of carbon dioxide in Earth's atmosphere. Triggering the onset of an ice age may require a sustained drop in global temperatures of only about 0.2–0.5 K, so astronomers are continuing to investigate a possible link between solar variability and changes in Earth's climate.

Questions and Problems

Test Your Understanding

1. The physical model of the Sun's interior has been confirmed by observations of
   a. neutrinos and seismic vibrations.
   b. sunspots and solar flares.
   c. neutrinos and positrons.
   d. sample returns from spacecraft.
   e. sunspots and seismic vibrations.
2. Place in order the following steps in the fusion of hydrogen into helium. If two or more steps happen simultaneously, use an equals sign (=).
   a. A positron is emitted.
   b. One gamma ray is emitted.
   c. Two hydrogen nuclei are emitted.
   d. Two ^{3}He collide and become ^{4}He.
   e. Two hydrogen nuclei collide and become ^{2}H.
   f. Two gamma rays are emitted.
   g. A neutrino is emitted.
   h. One deuterium nucleus and one hydrogen nucleus collide and become ^{3}He.
3. Sunspots, flares, prominences, and coronal mass ejections are all caused by
   a. magnetic activity on the Sun.
   b. electrical activity on the Sun.
   c. the interaction of the Sun's magnetic field and the interstellar medium.
   d. the interaction of the solar wind and Earth's magnetic field.
   e. the interaction of the solar wind and the Sun's magnetic field.
4. The structure of the Sun is determined by both the balance between the forces due to __________ and gravity and the balance between energy generation and energy __________.
   a. pressure; production
   b. pressure; loss
   c. ions; loss
   d. solar wind; production
5. In the proton-proton chain, four hydrogen nuclei are converted to a helium nucleus. This does not happen spontaneously on Earth because the process requires
   a. vast amounts of hydrogen.
   b. very high temperatures and densities.
   c. hydrostatic equilibrium.
   d. very strong magnetic fields.
6. The solar neutrino problem pointed to a fundamental gap in our knowledge of
   a. nuclear fusion.
   b. neutrinos.
   c. hydrostatic equilibrium.
   d. magnetic fields.
7. Sunspots change in number and location during the solar cycle. This phenomenon is connected to
   a. the rotation rate of the Sun.
   b. the temperature of the Sun.
   c. the magnetic field of the Sun.
   d. the tilt of the axis of the Sun.
8. Suppose an abnormally large amount of hydrogen suddenly burned in the core of the Sun. Which of the following would be observed first?
   a. The Sun would become brighter.
   b. The Sun would swell and become larger.
   c. The Sun would become bluer.
   d. The Sun would emit more neutrinos.
9. The solar corona has a temperature of 1 million to 2 million K; the photosphere has a temperature of only about 6000 K. Why isn't the corona much, much brighter than the photosphere?
   a. The magnetic field traps the light.
   b. The corona emits only X-rays.
   c. The photosphere is closer to us.
   d. The corona has a much lower density.
10. The Sun rotates once every 25 days relative to the stars. The Sun rotates once every 27 days relative to Earth. Why are these two numbers different?
    a. The stars are farther away.
    b. Earth is smaller.
    c. Earth moves in its orbit during this time.
    d. The Sun moves relative to the stars.
11. Place the following regions of the Sun in order of increasing radius.
    a. corona
    b. core
    c. radiative zone
    d. convective zone
    e. chromosphere
    f. photosphere
    g. a sunspot
12. Coronal mass ejections
    a. carry away 1 percent of the mass of the Sun each year.
    b. are caused by breaking magnetic fields.
    c. are always emitted in the direction of Earth.
    d. are unimportant to life on Earth.

13. As energy moves out from the Sun's core toward its surface, it first travels by __________, then by __________, and then by __________.
   a. radiation; conduction; radiation
   b. conduction; radiation; convection
   c. radiation; convection; radiation
   d. radiation; convection; conduction

14. Energy is produced primarily in the center of the Sun because
   a. the strong nuclear force is too weak elsewhere.
   b. that's where neutrinos are created.
   c. that's where most of the helium is.
   d. the outer parts have lower temperatures and densities.

15. The solar wind pushes on the magnetosphere of Earth, changing its shape, because
   a. the solar wind is so dense.
   b. the magnetosphere is so weak.
   c. the solar wind contains charged particles.
   d. the solar wind is so fast.

Thinking about the Concepts

16. Explain how hydrostatic equilibrium acts as a safety valve to keep the Sun at its constant size, temperature, and luminosity.

17. Two of the three atoms in a molecule of water (H_2O) are hydrogen. Why are Earth's oceans not fusing hydrogen into helium and setting Earth ablaze?

18. Why are neutrinos so difficult to detect?

19. Explain the proton-proton chain through which the Sun generates energy by converting hydrogen to helium.

20. On Earth, nuclear power plants use *fission* to generate electricity. In fission, a heavy element like uranium is broken into many atoms, where the total mass of the fragments is less than that of the original atom. Explain why fission could not be powering the Sun today.

21. If an abnormally large amount of hydrogen suddenly burned in the core of the Sun, what would happen to the rest of the Sun? Would the Sun change as seen from Earth?

22. Study the Process of Science Figure. If the follow-up experiments did not detect the other types of neutrinos, what would have been the next step for scientists at that point?

23. What is the solar neutrino problem, and how was it solved?

24. The Sun's visible "surface" is not a true surface, but a feature called the photosphere. Explain why the photosphere is not a true surface.

25. How are orbiting satellites and telescopes affected by the Sun?

26. Describe the solar corona. Under what circumstances can it be seen without special instruments?

27. In the proton-proton chain, the mass of four protons is slightly greater than the mass of a helium nucleus. Explain what happens to this "lost" mass.

28. What have sunspots revealed about the Sun's rotation?

29. Why are different parts of the Sun best studied at different wavelengths? Which parts are best studied from space?

30. Why is it important to study the interaction of the solar wind with the interstellar medium?

Applying the Concepts

31. In Figure 14.10, density and temperature are both graphed versus height.
   a. Is the height axis linear or logarithmic? How do you know?
   b. Is the density axis linear or logarithmic? How do you know?
   c. Is the temperature axis linear or logarithmic? How do you know?

32. Using the data in Figures 14.19b and c, present an argument that sunspots occur in regions of strong magnetic field.

33. Study Figure 14.17a and the graph in Figure 14.20.
   a. Estimate the fraction of the Sun's surface that is covered by the large sunspot group in the image. (Remember that you are seeing only one hemisphere of the Sun.)
   b. From the graph, estimate the average number of sunspots that occurs at solar maximum.
   c. On average, what fraction of the Sun could be covered by sunspots at solar maximum? Is this a large fraction?
   d. Compare your conclusion to the graph of irradiance in Figure 14.24. Does this graph make sense to you?

34. The Sun has a radius equal to about 2.3 light-seconds. Explain why a gamma ray produced in the Sun's core does not emerge from the Sun's surface 2.3 seconds later.

35. Assume that the Sun's mass is about 300,000 Earth masses and that its radius is about 100 times that of Earth. The density of Earth is about 5,500 kg/m^3.
   a. What is the average density of the Sun?
   b. How does this compare with the density of Earth? With the density of water?

36. The Sun shines by converting mass into energy according to $E = mc^2$. Show that if the Sun produces 3.85×10^{26} J of energy per second, it must convert 4.3 million metric tons (4.3×10^9 kg) of mass per second into energy.

37. Assume that the Sun has been producing energy at a constant rate over its lifetime of 4.6 billion years (1.4×10^{17} seconds).
    a. How much mass has it lost creating energy over its lifetime?
    b. The current mass of the Sun is 2×10^{30} kg. What fraction of its current mass has been converted into energy over the lifetime of the Sun?

38. Suppose our Sun was an A5 main-sequence star, with twice the mass and 12 times the luminosity of the Sun, a G2 star. How long would this A5 star burn hydrogen to helium? What would this mean for Earth?

39. Imagine that the source of energy in the interior of the Sun changed abruptly.
    a. How long would it take before a neutrino telescope detected the event?
    b. When would a visible-light telescope see evidence of the change?

40. On average, how long does it take particles in the solar wind to reach Earth from the Sun if they are traveling at an average speed of 400 km/s?

41. A sunspot appears only 70 percent as bright as the surrounding photosphere. The photosphere has a temperature of approximately 5780 K. What is the temperature of the sunspot?

42. The hydrogen bomb represents an effort to create a similar process to what takes place in the core of the Sun. The energy released by a 5-megaton hydrogen bomb is 2×10^{16} J.
    a. This textbook, *21st Century Astronomy*, has a mass of about 1.6 kg. If all of its mass were converted into energy, how many 5-megaton bombs would it take to equal that energy?
    b. How much mass did Earth lose each time a 5-megaton hydrogen bomb was exploded?

43. Verify the claim made at the start of this chapter that the Sun produces more energy per second than all the electric power plants on Earth could generate in a half-million years. Estimate or look up how many power plants there are on the planet and how much energy an average power plant produces. Be sure to account for different kinds of power; for example, coal, nuclear, wind.

44. Let's examine the reason that the Sun cannot power itself by chemical reactions. Using Working It Out 14.1 and the fact that an average chemical reaction between two atoms releases 1.6×10^{-19} J, estimate how long the Sun could emit energy at its current luminosity. Compare that estimate to the known age of Earth.

45. The Sun could get energy from gravitational contraction for a time period of $(GM_{\text{Sun}}/R_{\text{Sun}}L_{\text{Sun}})$. How long would the Sun last at its current luminosity? (Be careful with units!)

USING THE WEB

46. a. Go to *QUEST*'s "Journey into the Sun" Web page (http://science.kqed.org/quest/video/journey-into-the-sun) to watch a short video on the Solar Dynamics Observatory (SDO), launched in 2010. Why is studying the magnetic field of the Sun so important? What is new and different about this observatory? What is the "Music of the Sun"?
    b. Go to the SDO website (http://sdo.gsfc.nasa.gov). Under "Data," select "The Sun Now" and view the Sun at many wavelengths. What activity do you observe in the images at the location of any sunspots seen in the "HMI Intensitygram" images? (You can download a free SDO app by Astra to get real-time images on your mobile device.) Look at a recent news story from the SDO website. What was observed, and why is it newsworthy?
    c. Go to the *STEREO* mission's website (http://stereo.gsfc.nasa.gov). What is *STEREO*? Where are the spacecraft located? How does this configuration enable observations of the entire Sun at once? (You can download the app "3-D Sun" to get the latest images on your mobile device.)

47. a. What are the science goals of NASA's *Interface Region Imaging Spectrograph* (*IRIS*) mission (www.nasa.gov/mission_pages/iris/)? What has it discovered?
    b. An older space mission, *SOHO* (Solar and Heliospheric Observatory; http://sohowww.nascom.nasa.gov), was launched in 1995 by NASA and ESA. Click on "The Sun Now" to see today's images. The Extreme Ultraviolet imaging Telescope (EIT) images are in the far ultraviolet and show violent activity. How do these images differ from the ones of SDO in question 46b?
    c. Go to the Daniel K. Inouye Solar Telescope (DKIST) website (http://atst.nso.edu). This adaptive-optics telescope under construction on Haleakala, Maui, will be the largest solar telescope. Why is it important to study the magnetic field of the Sun? What are some of the advantages of studying the Sun from a ground-based telescope instead of a space-based telescope? What wavelengths does the DKIST observe? Why is Maui a good location? When is the telescope scheduled to be completed?

48. a. Go to the Space Weather website (http://spaceweather.com). Are there any solar flares today? What is the sunspot number? Is it about what you would expect for this year? (Click on "What is the sunspot number?" to see a current graph.) Are there any coronal holes today?
    b. Citizen science: Go to the website for Sunspotter (http://www.sunspotter.org/), a Zooniverse project that evaluates the complexity of sunspots and how they change over time. Zooniverse projects offer an opportunity for people to

contribute to science by analyzing pieces of data. Create an account for Zooniverse if you don't already have one (you will use it again in this course). Log in and click on "Science" and skim through the sections. What are the goals of this project? Why is it useful to have multiple people looking at these data? Click on "Classify" and analyze some sunspots. Save a screen shot for your homework.

c. Citizen science: Go to the Solar Stormwatch website (http://solarstormwatch.com), a Zooniverse project from the Royal Observatory in Greenwich, England. Create an account for Zooniverse if you don't already have one (you will use it again in this course). Log in and click on "Spot and Track Storms" and go through the Spot and Track training exercises. You are now ready to look at some real data. Click on an image to do the classification. Save a screen shot for your homework.

49. a. Go to the National Ignition Facility (NIF) website (https://lasers.llnl.gov/about/). Under "Science," click on "How to Make a Star." How are lasers used in experiments to develop controlled nuclear fusion on Earth? How does the fusion reaction here differ from that in the Sun?

b. An alternative approach is to fuse ^{3}He + ^{3}He instead of the hydrogen isotopes. But on Earth, ^{3}He is in limited supply. Helium-3 is in much greater abundance on the Moon, so some people propose setting up mining colonies on the Moon to extract ^{3}He for fusion reactions on Earth. Do a search on "helium 3 Moon." Which countries are talking about going to the Moon for this purpose? What is the timeline for when this might happen? What are the difficulties?

50. a. Go to http://voyager.jpl.nasa.gov/where/. Where are the *Voyager* spacecraft now? Has *Voyager 2* crossed into interstellar space?

b. Go to the website for the Interstellar Boundary Explorer (*IBEX*) (http://www.nasa.gov/mission_pages/ibex/.) What has *IBEX* learned about the solar wind and the interstellar medium?

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

EXPLORATION

The Proton-Proton Chain

digital.wwnorton.com/astro5

The proton-proton chain powers the Sun by fusing hydrogen into helium. This fusion process produces several different particles as byproducts, as well as energy. In this Exploration, we will explore the steps of the proton-proton chain in detail, with the intent of helping you keep them straight.

Visit the Student Site at the Digital Landing Page, and open the "Proton-Proton Animation" for this chapter.

Watch the animation all the way through once.

Play the animation again, pausing after the first collision. Two hydrogen nuclei (both positively charged) have collided to produce a new nucleus with only one positive charge.

1 Which particle carried away the other positive charge?

____

____

2 What is a neutrino? Did the neutrino enter the reaction or was the neutrino produced in the reaction?

____

____

Compare the interaction on the top with the interaction on the bottom.

3 Did the same reaction occur in each instance?

____

____

Resume playing the animation, pausing it after the second collision.

4 What two types of nuclei entered the collision? What type of nucleus resulted?

____

____

5 Was charge conserved in this reaction or was it necessary for a particle to carry charge away?

____

____

6 What is a gamma ray? Did the gamma ray enter the reaction or was it produced by the reaction?

____

____

Resume the animation again, and allow it to run to the end.

7 What nuclei enter the final collision? What nuclei are produced?

____

____

8 In chemistry, a catalyst facilitates the reaction but is not used up in the process. Do any nuclei act like catalysts in the proton-proton chain?

____

____

Make a table of inputs and outputs. Which of the particles in the final frame of the animation were inputs to the reaction? Which were outputs? Fill in your table with these inputs and outputs.

9 Which outputs are converted into energy that leaves the Sun as light?

____

____

10 Which outputs could become involved in another reaction immediately?

____

____

11 Which output is likely to stay in that form for a very long time?

____

____

15 The Interstellar Medium and Star Formation

The birth of a star—from a cloud of gas and dust to nuclear burning—is a process that can happen within tens of thousands of years (for the most massive stars) or require hundreds of millions of years (for the least massive). Astronomers have come to understand the process by observing many different stars at various stages of development. In this chapter, we will look at the interstellar environment from which stars both large and small form. Then we will focus our attention on the forming star—the *protostar*—and discuss how it becomes a star.

LEARNING GOALS

By the conclusion of the chapter, you should be able to:

LG 1 Describe the types and states of material that exist in the space between the stars and how this material is detected.

LG 2 Explain the conditions under which a cloud of gas can contract into a stellar system and the role that gravity and angular momentum play in the formation of stars and planets.

LG 3 List the steps in the evolution of a protostar.

LG 4 Describe the track of a protostar as it evolves to a main-sequence star on the Hertzsprung-Russell (H-R) diagram.

Stars are forming in this giant molecular cloud. ▶▶▶

How do stars form?

15.1 The Interstellar Medium Fills the Space between the Stars

The space between the stars is far from empty. It contains giant clouds of cool gas and dust, which can be observed in the visible part of the spectrum as they absorb the light from objects behind them, emit light from excited atoms, or reflect starlight from nearby stars. In turn, the space between these giant clouds is filled with tenuous gas. Together, all the material between the stars is known as the **interstellar medium**. Stars supply the energy that heats and stirs the interstellar medium. Warm gas heated by ultraviolet radiation from massive, hot stars pushes outward into its surroundings. Blast waves from dying stars sweep out vast, hot "cavities," piling up material in their path like snow in front of a snowplow. Interstellar clouds are both destroyed and formed by these violent events. Swept-up gas becomes the next generation of clouds. Hot bubbles of high-pressure gas crush molecular clouds, driving up their densities and sometimes triggering the formation of new generations of stars. Stars come from, and return much of their material to, the interstellar medium. In this section, we survey these varied states of the interstellar medium.

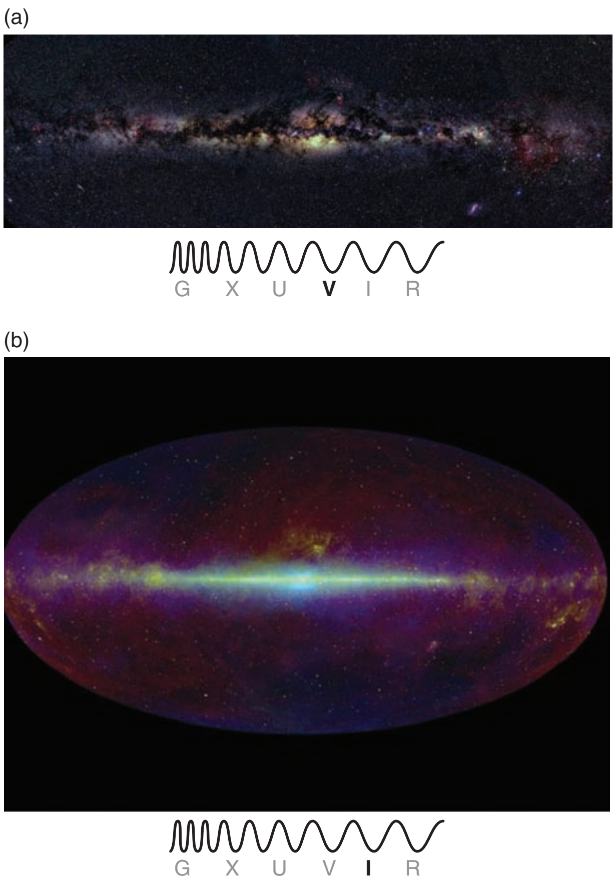


Figure 15.1 (a) This all-sky picture of the Milky Way was taken in visible light. The dark splotches blocking the view are dusty interstellar clouds. The center of this and all the other all-sky images in this chapter is the center of the Milky Way. (b) This all-sky picture was taken in the near infrared. Infrared radiation penetrates the interstellar dust, providing a clearer view of the stars in the disk of the Milky Way.

The Composition and Density of the Interstellar Medium

The Sun formed from the interstellar medium, so it is not too surprising that the chemical composition of the interstellar medium in the region of the Sun is similar to the chemical composition of the Sun (see Table 13.1). In the interstellar medium, hydrogen accounts for about 90 percent of the atomic nuclei, and the remaining 10 percent is almost all helium. The more massive elements account for only 0.1 percent of the atomic nuclei, or about 2 percent of the mass in the interstellar medium. Roughly 99 percent of that interstellar matter is gaseous, consisting of individual atoms or molecules moving about freely, as the molecules in the air do.

However, interstellar gas is far less dense than the air that you breathe. Each cubic centimeter (cm^3) of the air around you contains about 2.7×10^{19} molecules. A good vacuum pump on Earth can reduce this density down to about 10^{10} molecules/cm^3—approximately a billionth as dense. By comparison, the interstellar medium has an average density of about 0.1 atom/cm^3—one 100-billionth as dense as the vacuum pump can attain. Stated another way, there is about as much material in a column of air between your eye and the floor as there is interstellar gas in a column of the same diameter that stretches from the Solar System to the center of our galaxy 26,000 light-years away.

Interstellar Dust and Its Effects on Light

About 1 percent of the material in the interstellar medium is in the form of solid grains, called **interstellar dust**. Ranging in size from little more than large molecules up to particles about 300 nanometers (nm) across, these solid grains more closely resemble the particles of soot from a candle flame than the dust that collects on a windowsill. It would take several hundred "large" interstellar grains to span the thickness of a single human hair. Interstellar dust begins to form when materials such as iron, silicon, and carbon stick together to form grains in dense, relatively cool environments such as the outer atmospheres and "stellar winds" of cool, red giant stars or in dense material thrown into space by stellar explosions. Once these grains are in the interstellar medium, other atoms and molecules stick to them. This process is remarkably efficient: about half of the interstellar matter that is more massive than helium (1 percent of the total mass of the interstellar medium) is found in interstellar grains.

Because interstellar dust is extremely effective at blocking and diverting light, the view of distant objects is affected even by the low-density interstellar material. Recall the comparison of the air column between you and the floor and the column of interstellar space stretching to the center of the galaxy. If the interstellar column were compressed to the same density as air, it would be so dirty that it would be difficult to see your hand 10 centimeters (cm) in front of your face. Go out on a dark summer night in the Northern Hemisphere (or on a dark winter night in the Southern Hemisphere) and look closely at the Milky Way, visible as a faint band of diffuse light running through the constellation Sagittarius. You will see a dark "lane" running roughly down the middle of this bright band, splitting it in two. This dark band is a vast expanse of interstellar dust and blocks astronomers' views of distant stars.

When interstellar dust gets in the way of radiation from distant objects, the effect is called **interstellar extinction**. Not all electromagnetic radiation is affected equally by interstellar extinction. **Figure 15.1** shows two images of the Milky Way Galaxy: one taken in visible light, the other taken in the infrared (IR). The dark clouds that block the shorter-wavelength visible light (Figure 15.1a) seem to have vanished in the longer-wavelength IR image (Figure 15.1b), enabling observations through the clouds to the center of the galaxy and beyond. These two images are *all-sky images*: they portray the entire sky surrounding Earth. They have been oriented so that the disk of the Milky Way, in which the Solar System is embedded, runs horizontally across the center of the image.

To understand why short-wavelength radiation is obscured by dust while long-wavelength radiation is not, think about a different kind of wave—waves on the surface of the ocean, as shown in **Figure 15.2a**. Imagine you are on the ocean in a boat in a strong swell. If the waves are much bigger than your small boat, the swell causes you to bob gently up and down. But that is about all: there is no other interaction between the waves and the small boat, and no energy is lost from the wave to the boat. The situation is quite different if the waves are closer in size to your boat. To picture this, imagine waves roughly half the size of your boat. Now, the front of the boat may be on a wave crest while the back of the boat is in a trough or vice versa. The boat tips wildly back and forth as the waves go by. If the size of the boat and the wavelength of the waves are the right match, even fairly modest waves will rock the boat. You might have noticed this if you were ever in a canoe or rowboat when the wake from a speedboat came by. Now imagine viewing these two situations from the perspective of the wave. The wave is hardly affected when it is much bigger than the boat, but it is strongly affected when it is comparable in size or small compared to the boat. The energy to drive the wild motions of the boat comes from the wave, so the wave loses energy.

The interaction of electromagnetic waves with matter is more involved than that of a boat rocking on the ocean, but the same basic idea often applies, as shown in Figure 15.2b. Tiny interstellar dust grains effectively block the transmission of ultraviolet light and blue light, which have wavelengths comparable to or smaller than the typical size of dust grains. In contrast, longer-wavelength infrared and radio radiation does not interact strongly with the tiny interstellar dust grains. Therefore, at visible and ultraviolet wavelengths, most of the Milky Way is hidden from view by dust, but in the infrared and radio portions of the spectrum, we get a far more complete view.

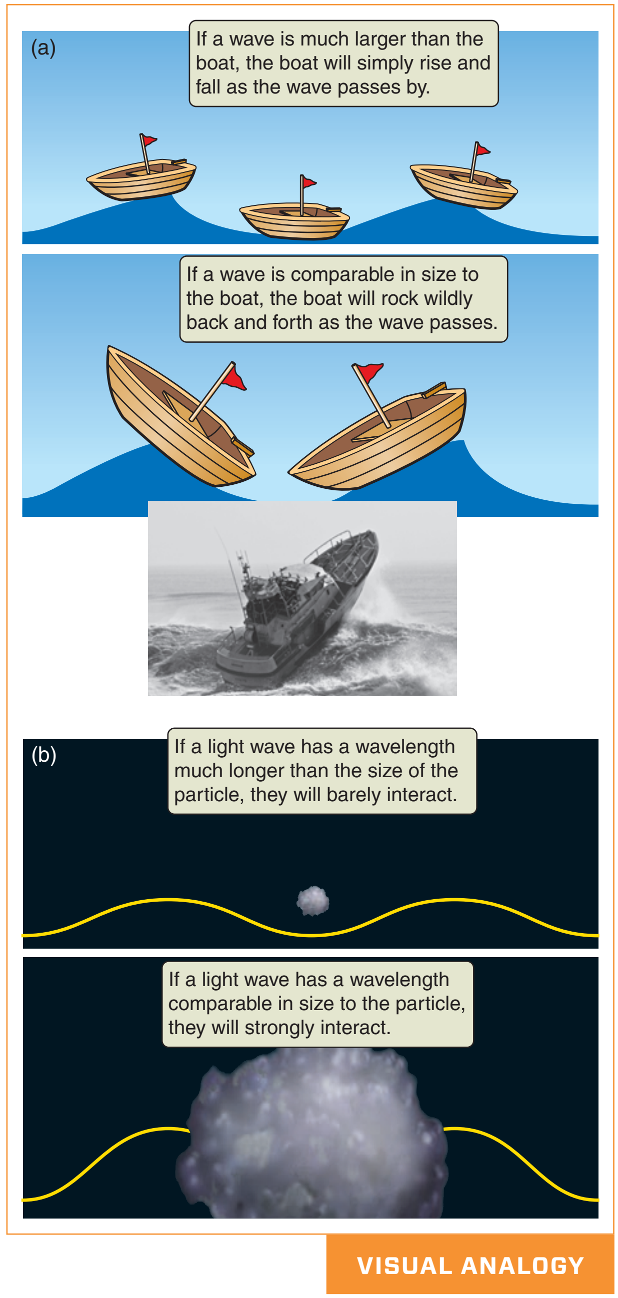


Figure 15.2 (a) Just as boats interact most strongly with ocean waves that are similar in size, (b) particles interact most strongly with wavelengths of light of similar size.

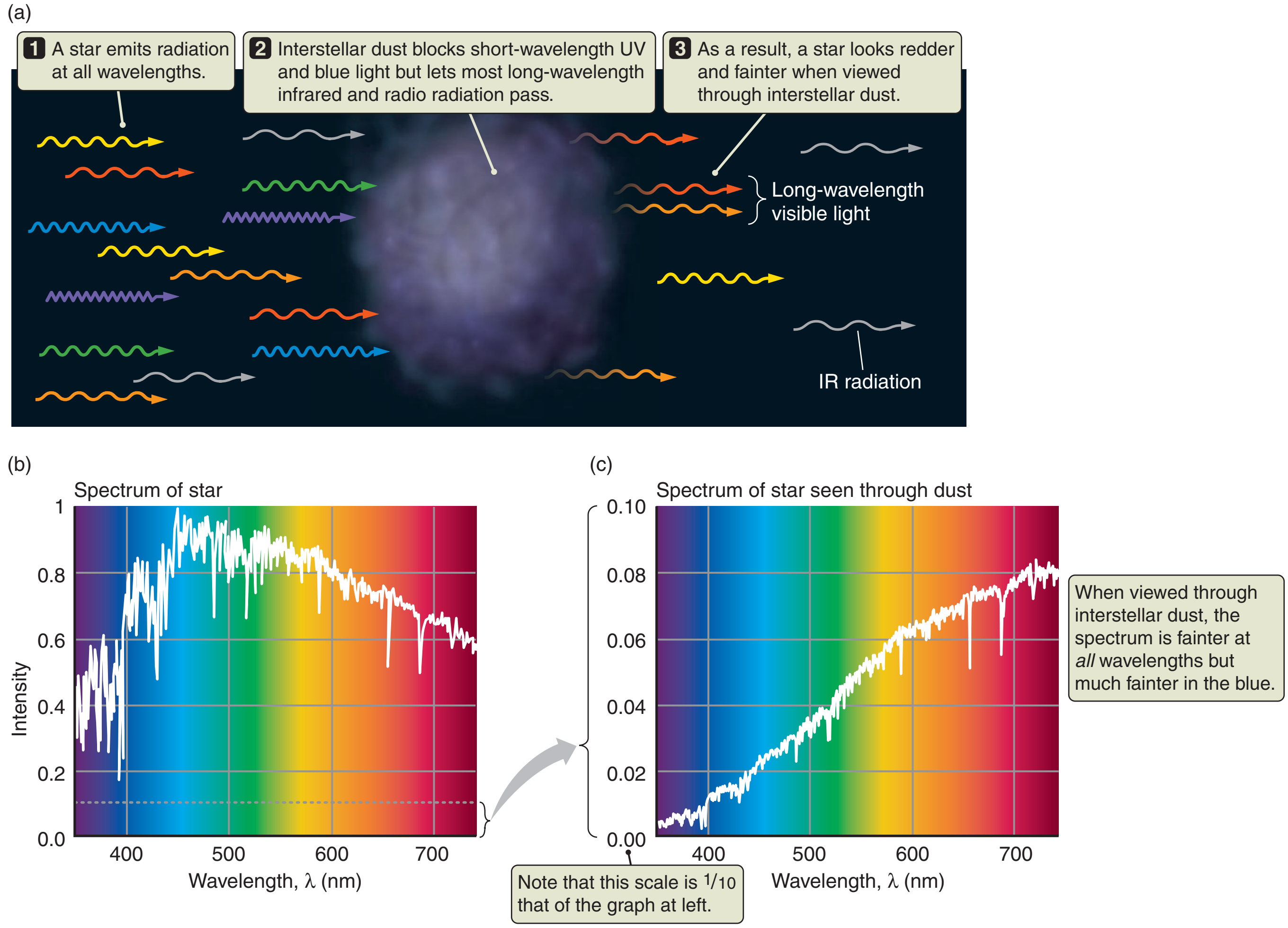


Figure 15.3 (a) The wavelengths of ultraviolet and blue light are close to the size of interstellar grains, so the grains effectively block this light. Grains are less effective at blocking longer-wavelength light. As a result, the spectrum of a star (b) when seen through an interstellar cloud (c) appears fainter and redder.

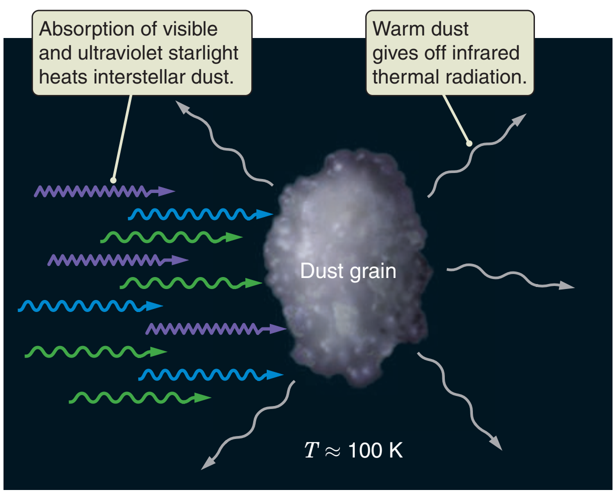


Figure 15.4 The temperature of interstellar grains is determined by the equilibrium between absorbed and emitted radiation.

As shown in **Figure 15.3**, extinction occurs at all wavelengths, causing an object viewed through dust to be fainter than it would be otherwise. However, extinction affects short-wavelength blue light more than it affects long-wavelength red light. This makes the object appear less blue than it really is. This effect of dust, of removing the blue light to cause an object to be more red, is called **reddening**. The presence of dust significantly affects the spectrum of an object, as shown in Figures 15.3b and c. Correcting for the presence of dust can be one of the most difficult parts of interpreting astronomical observations, often adding to uncertainty in the measurement of an object's properties.

Interstellar extinction is less of a concern at infrared wavelengths, but dust still plays an important role in infrared observations. Like any other solid object, a large grain of dust glows at a wavelength determined by its temperature. In Chapter 5, we discussed how the equilibrium between absorbed sunlight and emitted thermal radiation determines the temperatures of the terrestrial planets (see Figure 5.20). As **Figure 15.4** illustrates, a similar equilibrium is at work in interstellar space, where dust is heated both by starlight and by the gas in which it is immersed to temperatures of tens to hundreds of kelvins. At a temperature of 100 kelvins (K), Wien's law says that dust will glow most strongly at a wavelength

15.1 Working It Out Dust Glows in the Infrared

The temperature of interstellar dust can be found from its spectrum. Wien's law, discussed in Chapter 5, relates the temperature of an object to the peak wavelength (λ_{peak}) of its emitted radiation. For warm dust at a temperature of 100 K (recall that 1 μm = 10^{-6} meter = 1,000 nm):

$$\lambda_{peak} = \frac{2{,}900\ \mu\text{m K}}{T} = \frac{2{,}900\ \mu\text{m K}}{100\ \text{K}} = 29\ \mu\text{m}$$

For cooler dust, at a temperature of 10 K:

$$\lambda_{peak} = \frac{2{,}900\ \mu\text{m K}}{T} = \frac{2{,}900\ \mu\text{m K}}{10\ \text{K}} = 290\ \mu\text{m}$$

The temperature and the peak wavelength are inversely proportional, so if the temperature drops, the peak wavelength gets longer. For the temperatures that are common for dust in the interstellar medium, the peak wavelength is in the far-infrared part of the electromagnetic spectrum.

of 29 microns (μm), whereas cooler dust—at a temperature of 10 K—glows most strongly at a longer wavelength (**Working It Out 15.1**). Much of the light in infrared observations is thermal radiation from this dust. **Figure 15.5a**, from NASA's Wide-field Infrared Survey Explorer (WISE), shows the sky at combined wavelengths of 3.4, 12, and 22 μm. In Figure 15.5b, a far-infrared image of the sky at 100 μm from the Infrared Astronomical Satellite (IRAS) telescope is combined with a microwave image from the Cosmic Background Explorer (COBE) telescope to show the Milky Way's dark clouds glowing brilliantly in infrared radiation from dust.

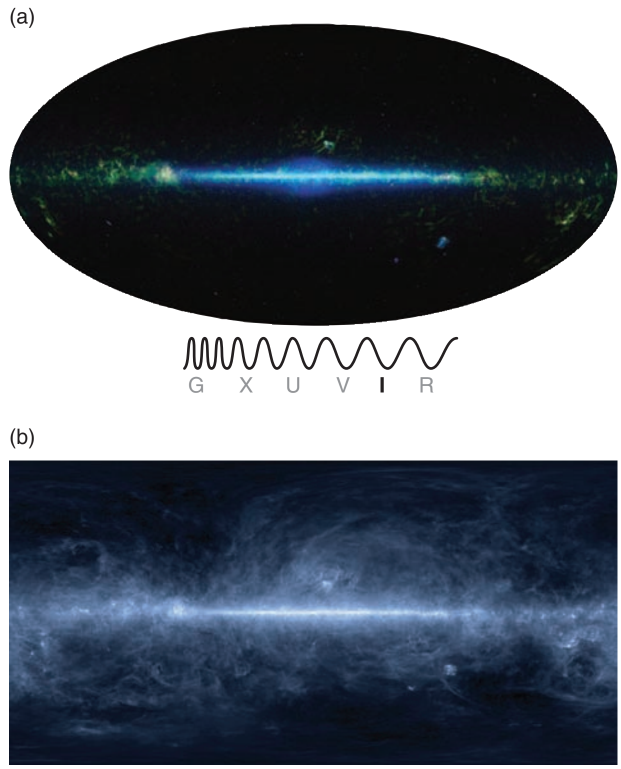


Figure 15.5 (a) This WISE infrared image shows the plane of the Milky Way at combined wavelengths of 3.4, 12, and 22 μm. (b) This all-sky image, in the far-infrared wavelength of 100 μm (from IRAS) combined with a microwave image from COBE, shows dust throughout the galaxy.

Temperatures and Densities of Interstellar Gas

About half of the gas and dust that fills interstellar space is concentrated in dense regions called **interstellar clouds**, in which the interstellar gas is more concentrated than in surrounding regions. These clouds fill only about 2 percent of the volume of interstellar space. The other half of the interstellar gas and dust is spread out through 98 percent of the volume of interstellar space and is called **intercloud gas**.

The properties of intercloud gas vary from place to place (**Table 15.1**). About half of the volume of interstellar space is filled with an intercloud gas that is extremely hot, heated primarily by the energy of tremendous stellar explosions

TABLE 15.1 Typical Properties of Components of the Interstellar Medium

Component	Temperature (K)	Number Density (atoms/cm³)	Size of Cube per Gram* (km)	State of Hydrogen
Hot intercloud gas	~1 million	~0.005	~8,000	Ionized
Warm intercloud gas	~8000	0.01–1	~800	Ionized or neutral
Cold intercloud gas	~100	1–100	~80	Neutral
Interstellar clouds	~10	100–1,000	~8	Molecular or neutral

*This is the length of one side of the cube of space you would need to search to find 1 gram of the material. It is another way of thinking about density.

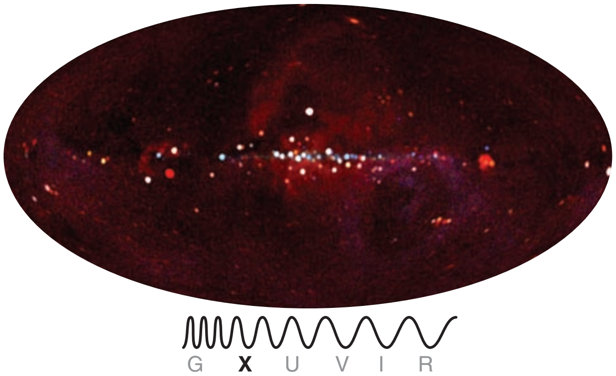


Figure 15.6 Bright spots in this image are distant X-ray sources, including objects such as bubbles of very hot, high-pressure gas surrounding the sites of recent explosions of supernovae.

called *supernovae* (**Figure 15.6**). This hot intercloud gas has temperatures like those found in the cores of stars—in the millions of kelvins. Because the temperature is so high, the atoms in the gas are moving very rapidly. But the gas also has an extremely low density. Typically, you would have to search a liter (1,000 cm^3) or more of hot intercloud gas to find a single atom. If you were adrift in an expanse of this million-kelvin intercloud gas, it would do little to keep you warm. Even though the atoms are moving quickly and each one would impart a lot of energy if it collided with you, there are so few atoms that collisions would be rare. You would radiate energy away and cool off much faster than the gas around you could replace the lost energy.

The Solar System is passing through a bubble of this hot intercloud gas that has a density of about 0.005 hydrogen atom/cm^3 and is at least 650 light-years across. This may be the remnant of the hot bubble produced by a supernova explosion 300,000 years ago. Like the million-kelvin gas in the corona of the Sun, hot intercloud gas glows faintly in the energetic X-ray portion of the electromagnetic spectrum. Orbiting X-ray telescopes observe the entire sky aglow with faint X-rays coming from our local bubble of million-kelvin gas.

Not all intercloud gas is as hot as that of the local bubble. Most other intercloud gas is "warm" and has a temperature of about 8000 K and a density ranging from about 0.01 to 1 atom/cm^3. About half of the volume of warm intercloud gas is kept ionized by starlight. Ultraviolet starlight with wavelengths shorter than 91.2 nm has enough energy to **ionize** hydrogen, which means the electron has been stripped away from the nucleus. But if there is a large enough expanse of warm intercloud gas, then the outermost atoms shield the neutral gas nearer the center by "using up" all the ultraviolet light, much as Earth's ozone layer shields the surface of Earth from harmful ultraviolet radiation from the Sun. This shielding leaves much gas in an unionized state.

One way to look for interstellar gas, including both warm and hot interstellar gas, is to study the spectra of distant stars. Most commonly, this gas is observed in absorption lines in the spectra of distant stars. Atoms in the gas absorb starlight at particular wavelengths, indicating the temperature, density, and chemical composition of the gas. Intercloud gas can also sometimes produce emission lines in regions of warm, ionized gas, for example in supernova remnants, where protons and electrons constantly recombine into hydrogen atoms. When a proton and an electron combine to form a neutral hydrogen atom, energy must be given up in the form of electromagnetic radiation. Typically, the resulting hydrogen atom is left in an excited state (see Chapter 5). The atom then drops down to lower and lower energy states, emitting a photon at each step, so warm, ionized interstellar gas glows in emission lines characteristic of hydrogen. Usually, the strongest emission line given off by warm interstellar gas in the visible part of the spectrum is the Hα (hydrogen alpha) line, which is seen in the red part of the spectrum at a wavelength of 656.3 nm. Other elements undergo a similar process.

The faint, diffuse emission in **Figure 15.7** comes mostly from warm (about 8000 K), ionized intercloud gas, glowing in Hα. The bright spots are called **H II regions** ("H two") because the hydrogen atoms are ionized; that is, they are in the second state after neutral. In H II regions, intense ultraviolet radiation from massive, hot, luminous O and B stars is able to ionize even relatively dense interstellar clouds, causing Hα emission in a rough sphere around each star. O stars live only a few million years, so they usually do not move very far from where they formed. The glowing clouds seen as H II regions are the very clouds from which these stars were born and indicate active star formation.

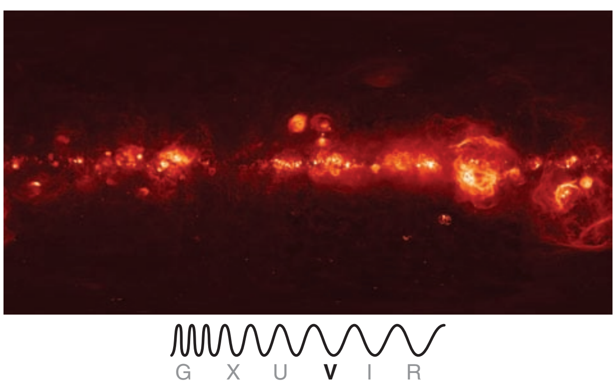


Figure 15.7 Warm interstellar gas (about 8000 K) glows in the Hα line of hydrogen. This image of the Hα emission from much of the northern sky reveals the complex structure of the interstellar medium.

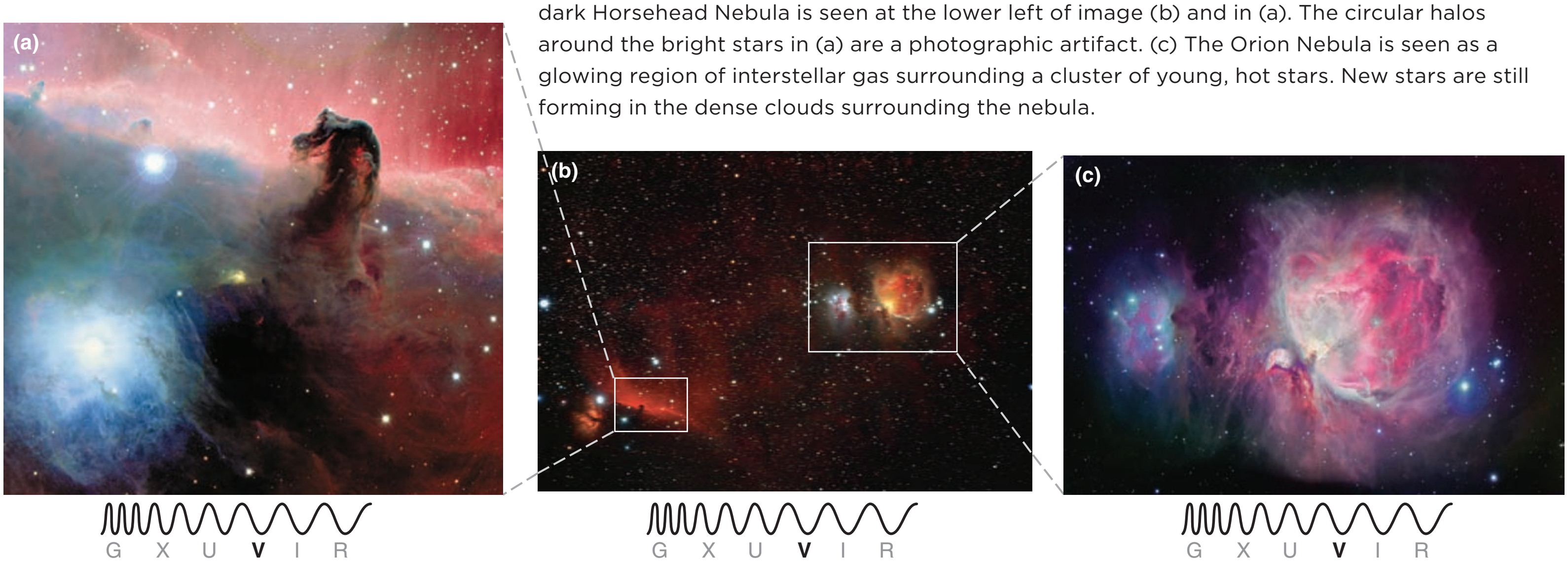


Figure 15.8 The Orion Nebula is only a small part of the larger Orion star-forming region. The dark Horsehead Nebula is seen at the lower left of image (b) and in (a). The circular halos around the bright stars in (a) are a photographic artifact. (c) The Orion Nebula is seen as a glowing region of interstellar gas surrounding a cluster of young, hot stars. New stars are still forming in the dense clouds surrounding the nebula.

One of the closest H II regions to the Sun is the Orion Nebula, located 1,340 light-years from the Sun in the constellation Orion (**Figure 15.8**). Almost all of the ultraviolet light that powers this nebula comes from a single hot star, and only a few hundred stars are forming in its immediate vicinity. In contrast, a dense star cluster containing thousands of hot, luminous stars powers a giant H II region called 30 Doradus, located in the Large Magellanic Cloud, a small companion galaxy to the Milky Way 160,000 light-years away. If 30 Doradus were as close as the Orion Nebula, it would be bright enough in the nighttime sky to cast shadows.

Warm, *neutral* hydrogen gas gives off radiation in a different way from warm, *ionized* hydrogen. Many subatomic particles, including protons and electrons, have a property called *spin* that causes them to behave as though each particle has a bar magnet, with a north and a south pole, built into it. As demonstrated in **Figure 15.9**, a hydrogen atom can exist in only two configurations: either the magnetic "poles" of the proton and electron point in opposite directions or they are aligned. These configurations have different energies. When the two "magnets" point in the same direction, the atom has slightly less energy then when they point in the opposite direction. If left undisturbed long enough, a hydrogen atom in the higher energy state will spontaneously jump to the lower energy state, emitting a photon in the process. The energy difference between the two magnetic spin states of a hydrogen atom is extremely small, so the emitted photon has a wavelength of 21 cm, in the radio region of the spectrum. Later, interactions between atoms in the gas will bump the hydrogen atoms back to the higher energy state, refreshing the supply of atoms that can produce the 21-cm line.

The tendency for hydrogen atoms to emit 21-cm radiation is extremely weak. On average, you would have to wait about 11 million years for a hydrogen atom in the higher energy state to jump spontaneously to the lower energy state and give off a photon. But there is a lot of hydrogen in the universe, so at any given time, some atoms are making this transition. In **Figure 15.10**, the sky is aglow with 21-cm radiation from neutral hydrogen. Because of its long wavelength, 21-cm radiation freely penetrates dust in the interstellar medium, enabling astronomers to see neutral hydrogen throughout the galaxy, while measurements of the Doppler shift of

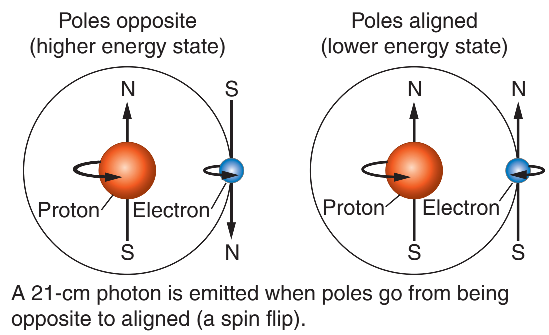


Figure 15.9 There is a slight difference in energy when the poles of the proton and electron are aligned compared to when they are opposite. This energy difference corresponds to a photon of 21 cm.

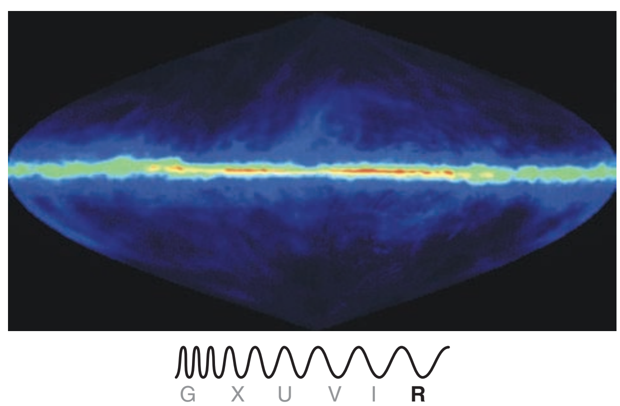


Figure 15.10 This radio image of the sky shows the distribution of neutral hydrogen gas throughout the galaxy. Red indicates directions of the highest hydrogen density, and blue and black show areas with little hydrogen. Radio waves penetrate interstellar dust, and probe of the structure of the galaxy.

the line indicate how fast the emitting gas is moving toward us or away from us. These two attributes make the 21-cm line of neutral hydrogen important for understanding the structure of our galaxy (**Process of Science Figure**).

Regions of Cool, Dense Gas

Most interstellar clouds are composed primarily of isolated neutral hydrogen atoms and are much cooler and denser than the warm intercloud gas. They have temperatures of about 100 K and densities in the range of about 1–100 atoms/cm^3. On average, atoms in interstellar clouds are moving around with random velocities of approximately 20 kilometers per second (km/s). The fastest motions of interstellar material are those of very hot gas and are measured in thousands of kilometers per second.

On Earth, it is uncommon to find atoms in isolation; most atoms are in molecules. For example, in Earth's atmosphere, only nonreactive gases such as argon are typically found in their atomic form. In most of interstellar space, however, including most interstellar clouds, molecules do not survive long. If interstellar gas is too hot, then any molecules that do exist soon collide with other molecules or atoms that have enough energy to break the molecules apart. The temperature in a neutral hydrogen cloud may be low enough for some molecules to survive, but photons of starlight with enough energy to break molecules apart can penetrate neutral hydrogen clouds. The hearts of the densest of interstellar clouds are known as **molecular clouds** because in these regions, dust effectively blocks even relatively low-energy photons, and so molecules persist.

Molecular cloud masses range from a few solar masses to 10 million solar masses. The smallest molecular clouds may be less than half a light-year across; the largest may be more than a thousand light-years in size. **Giant molecular clouds** are typically about 100–200 light-years across and have masses a few hundred thousand times that of the Sun. The Milky Way Galaxy contains several thousand giant molecular clouds and a much larger number of smaller ones. Molecular clouds fill only about 0.1 percent of interstellar space. These clouds may be rare, but they are extremely important because they are the cradles of star formation.

In images such as those of **Figure 15.11**, the dust in molecular clouds causes a silhouette in visible light against a background of stars. Infrared radiation passes through this dust, however, so we can see through the dust to sources inside and behind the cloud. Inside such clouds it is dark and usually very cold, with a typical temperature of only about 10 K. Most of these clouds have densities of about 100–1,000 molecules/cm^3, but densities as high as 10^{10} molecules/cm^3 have been observed. Even at 10^{10} molecules/cm^3, this gas is still less than a billionth as dense as the air around you, making it an extremely good vacuum by terrestrial standards. In this cold, relatively dense environment, atoms combine to form a wide variety of molecules.

By far the most common component of molecular clouds is molecular hydrogen. Molecular hydrogen (H_2) consists of two hydrogen atoms and is the smallest possible molecule. Molecules radiate as they make transitions between energy states: these states are determined by the way the molecules rotate or vibrate, for example. Molecular emission lines are useful in the same way that atomic emission lines are useful. Each type of molecule is unique in its properties, and thus unique in its energy states. The wavelengths of emission lines from molecules are an unmistakable fingerprint of the kinds of molecules responsible for them.

(a)

(b)

Figure 15.11 In visible light, interstellar molecular clouds are seen in silhouette against a background of stars and glowing gas. (a) Light from background stars is blocked by dust and gas in nearby Barnard 68, a dense, dark molecular cloud. (b) Infrared wavelengths can penetrate much of this gas and dust, as seen in this false-color image of Barnard 68.

Process of Science

ALL BRANCHES OF SCIENCE ARE INTERCONNECTED

Studies of the natural world on the smallest and largest scales inform one another.

Atomic physics: Scientists studying atoms and quantum mechanics—the underlying principles that govern the behavior of atoms—find that the electron in the hydrogen atom makes a rare transition, releasing radiation with a wavelength of 21 cm. This is so rarely observed on Earth that it is known as a *forbidden transition*.

Radio astronomy: Because space is so large, there are vast numbers of hydrogen atoms along the line of sight in every direction. There are so many that even exceptionally rare events happen often enough to be detectable. The forbidden transition is commonly observed by radio astronomers and maps out the neutral hydrogen in the Milky Way.

Radio astronomers use atomic physics to understand the behavior of the entire galaxy, which is examined at a scale 10^{30} times larger.

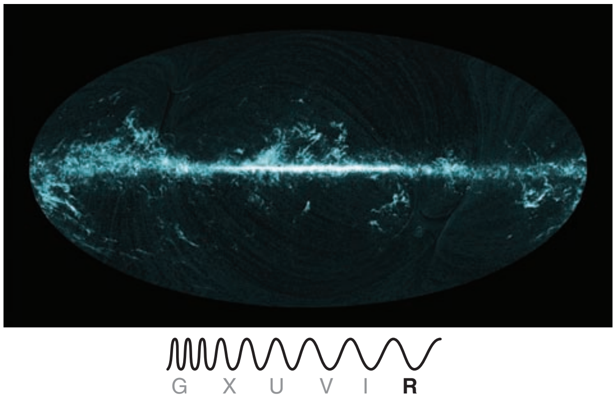


Figure 15.12 This all-sky image from the Planck observatory shows the distribution of carbon monoxide (CO), which traces molecular clouds where stars are born.

Because some of these transitions are in the radio or infrared portions of the spectrum, these molecules can be detected even deep inside a molecular cloud.

In addition to molecular hydrogen, approximately 150 other molecules have been observed in interstellar space. These molecules range from very simple structures such as carbon monoxide (CO), to complex organic compounds such as methanol (CH_3OH) and acetone $(CH_3)_2CO$, to some with large carbon chains. Very large carbon molecules, made of hundreds of individual atoms, bridge the gap between large interstellar molecules and small interstellar grains. Visible light cannot escape from the molecular clouds where interstellar molecules are concentrated. The radio waves emitted by molecules, however, are unaffected by interstellar dust, so they escape easily from dark molecular clouds. Observations of molecular lines reveal the innermost workings of the densest and most opaque interstellar clouds. Among the more important molecules is CO (**Figure 15.12**). The ratio of CO to H_2 is relatively constant, as far as has been tested, and interstellar carbon monoxide observed from Earth is often used to estimate the amounts and distribution of interstellar neutral molecular hydrogen, which is more difficult to observe directly.

CHECK YOUR UNDERSTANDING 15.1

When radiation from an object passes through the interstellar medium: (a) the object appears dimmer; (b) the object appears bluer; (c) the object appears bluer *and* dimmer; (d) the object appears redder *and* dimmer.

15.2 Molecular Clouds Are the Cradles of Star Formation

Stars—and planets—form from large clouds of dust and gas in the interstellar medium. This star formation is triggered by nearby energetic events such as the explosion of a dying massive star. In this section, we explore the first steps of this process as a cloud begins to contract and fragment to form stars.

Self-Gravity in the Molecular Cloud

Recall that in our discussion of the Sun in Chapter 14, we referred to the balance between gravity and pressure in a stable object as *hydrostatic equilibrium*. Interstellar clouds are not always in hydrostatic equilibrium—in most interstellar clouds, internal pressure is much stronger than the gravity that holds a cloud together (the "self-gravity"). Because gravity follows an inverse square law, the more spread out an object's mass, the weaker its self-gravity. In most interstellar clouds, the internal gas pressure pushing out is much stronger than self-gravity, so a cloud should expand. But the much hotter gas surrounding the clouds also exerts a pressure inward on a cloud. This external hot gas helps hold a cloud together and often provides a trigger for collapse.

If a cloud is massive enough and dense enough (or becomes so after a triggering event), self-gravity becomes important. As shown in **Figure 15.13**, each part of the cloud feels a gravitational attraction from every other part of the cloud. The sum of all these forces acting on a particular parcel of gas will always point toward the center of the cloud's mass. This is the *net force* and indicates the direction in which the parcel will begin to move. In massive, dense clouds, self-gravity

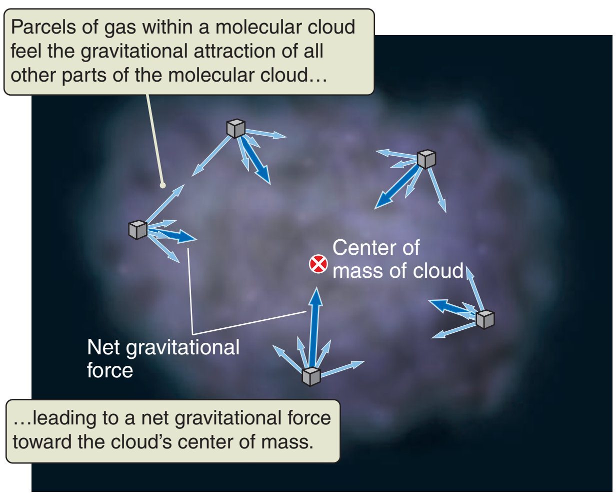


Figure 15.13 Self-gravity causes a molecular cloud to collapse, drawing parcels of gas toward a single point inside the cloud. The lighter blue arrows are examples of forces on the parcel due to other parcels of gas. The darker blue arrows show the sum of all these forces. This net force always points toward the center of mass of the cloud.

is much greater than pressure, so the clouds collapse under their own weight, beginning a chain of events that will form a new generation of stars.

If self-gravity in a molecular cloud is much greater than internal pressure, gravity should win outright, and the cloud should rapidly collapse toward its center. In practice, the process goes very slowly because several other effects stand in the way of the collapse. One effect that slows the collapse of a cloud is conservation of angular momentum, described in Chapter 7. Other effects that slow the collapse are turbulence, the effects of magnetic fields, and thermal pressure. Even though these effects may slow the collapse of a molecular cloud, in the end gravity will predominate. One part of the cloud can lose angular momentum to another part of the cloud, allowing the part of the cloud with less angular momentum to collapse further. Neutral matter crosses magnetic field lines, gradually increasing the gravitational pull toward the center, until the force on the charged particles is large enough to drag the magnetic field toward the center as well. Turbulence ultimately fades away. The details of these processes are complex and are the subject of much current research. For our purpose, the important point is that effects that prevent the collapse of a molecular cloud are temporary, and gravity is persistent. As the forces that oppose the cloud's self-gravity gradually fade, the cloud slowly shrinks.

Astronomy in Action: Angular Momentum

AstroTour: Star Formation

Molecular Clouds Fragment as They Collapse

Molecular clouds are clumpy. Some regions within the cloud are denser and collapse more rapidly than surrounding regions. As these regions collapse, their self-gravity becomes stronger because they are more compact, so they collapse even faster. **Figure 15.14** shows the process of collapse in a molecular cloud. Slight variations in the density of the cloud grow to become very dense concentrations of gas. Instead of collapsing into a single object, the molecular cloud fragments into very dense **molecular-cloud cores**. A single molecular cloud may form hundreds or thousands of molecular-cloud cores, each of which is typically a few light-months in size. Some of these dense cores will eventually form stars.

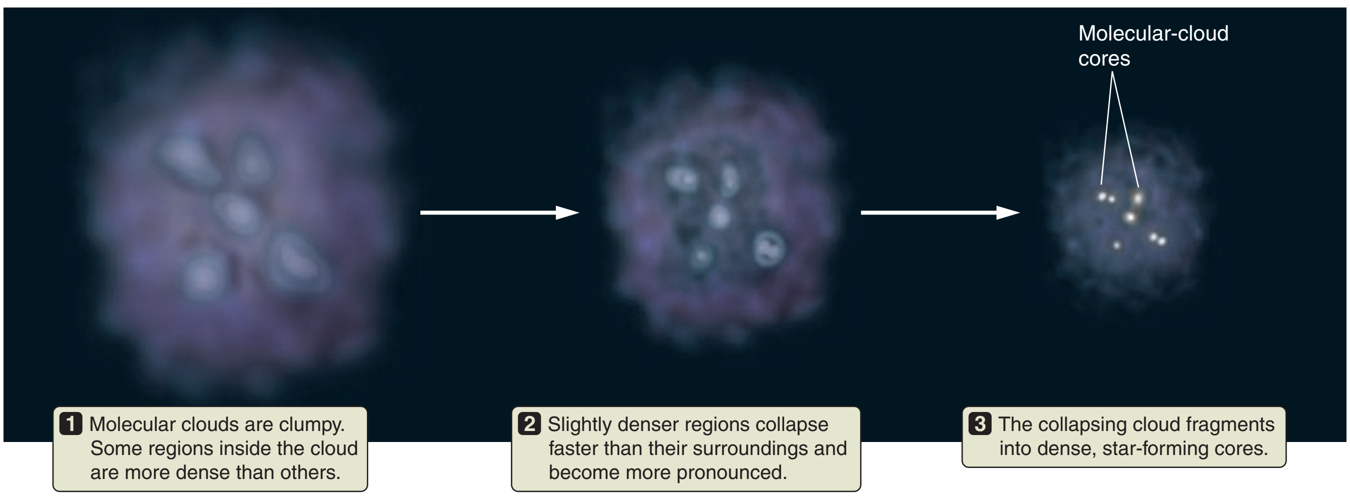


Figure 15.14 When a molecular cloud collapses, denser regions within the cloud collapse more rapidly than less dense regions. As this process continues, the cloud fragments into a number of very dense molecular-cloud cores that are embedded within the large cloud. These cloud cores may go on to form stars.

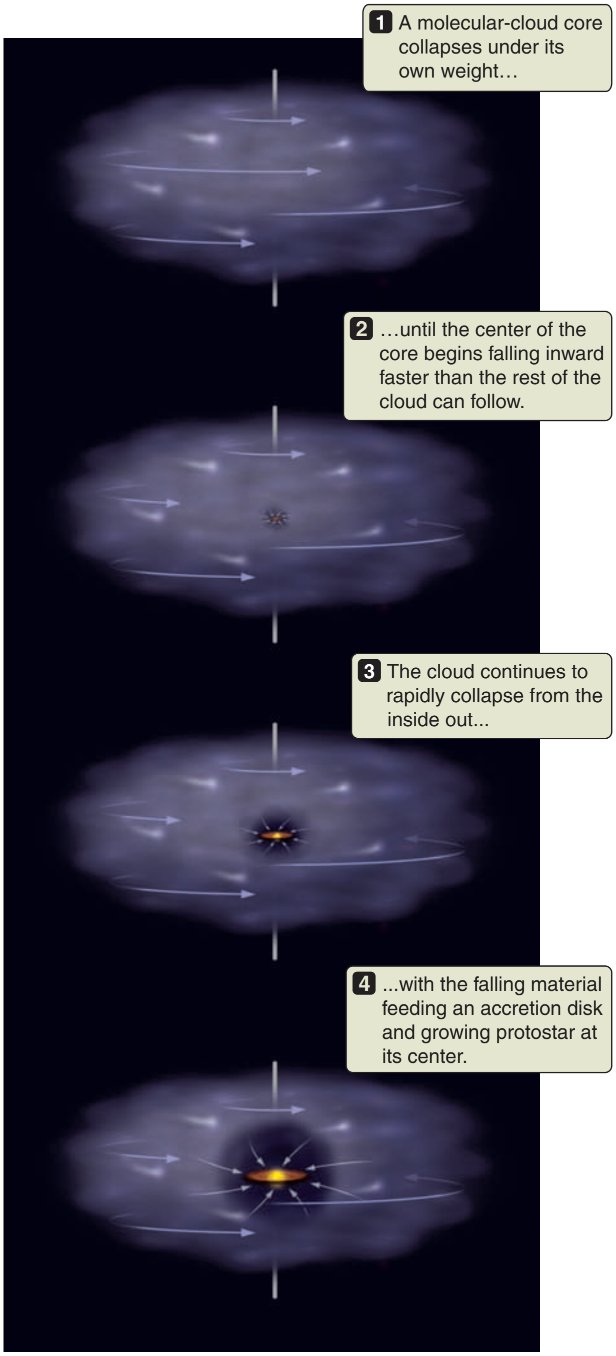


Figure 15.15 When a molecular-cloud core gets very dense, it collapses from the inside out. Conservation of angular momentum causes the infalling material to form an accretion disk that feeds the growing protostar.

As a molecular-cloud core collapses, the gravitational forces grow stronger still, because the force of gravity is inversely proportional to the square of the radius. Suppose a cloud is 4 light-years across. When the cloud has collapsed to 2 light-years across, the different parts of the cloud are, on average, only half as far apart as when the collapse started. As a result, the gravitational attraction the parts of the cloud feel toward each other is 4 times stronger. When the cloud is one-fourth as large as it was at the beginning of the collapse, the force of gravity is 16 times stronger. As a core collapses, the inward force of gravity increases; as gravity increases, the collapse speeds up; as the collapse speeds up, the gravitational force increases even faster.

Eventually, gravity is able to overwhelm the opposing forces due to pressure, magnetic fields, and turbulence. This happens first near the center of the cloud core where the cloud material is most strongly concentrated. The pressure from the central part of the cloud core supports the weight of the layers above it. The inner parts of the cloud core start to fall rapidly inward, removing support from the more distant parts of the cloud. Without the support of that inner material, the more distant material begins to fall freely toward the center. The cloud core collapses from the inside out, as shown in **Figure 15.15**.

CHECK YOUR UNDERSTANDING 15.2

Molecular clouds fragment as they collapse because: (a) the rotation of the cloud throws some mass to the outer regions; (b) the density increases fastest in the center of the cloud; (c) density variations from place to place grow larger as the cloud collapses; (d) the interstellar wind is stronger in some places than others.

15.3 Formation and Evolution of Protostars

Material from the collapsing molecular-cloud core falls inward. Because of conservation of angular momentum, this material accumulates in a flat, rotating accretion disk. Most of this material eventually finds its way inward to the center of the disk. The object that is forming there, on its way to becoming a star, is called a **protostar**. In this section, we follow the evolution of this object on the way to becoming a star.

A Protostar Forms

As the cloud collapses, gravitational energy is converted to thermal energy. Particles are pulled toward the center by gravity. As they fall, they move faster and faster. As they become more densely packed, they begin to crash into each other, causing random motions and raising the temperature of the core. These random motions of particles are collectively known as the *thermal energy*. When the particles are hotter, they are moving faster, and the thermal energy is higher. In the collapsing protostar, the thermal energy comes from the gravitational energy that was stored in the cloud when the particles were far apart. The collapse of the cloud converts gravitational energy to thermal energy, and the gas in the outer layers of the protostar is heated to a temperature of thousands of kelvins, causing the protostar to shine.

Because of this accumulation of thermal energy, the core gradually reaches more than 1 million K. At this temperature, deuterium can fuse with hydrogen to create helium-3. (Notice that this reaction occurs at a lower temperature than that at which hydrogen burning occurs.) Once deuterium burning begins, it drives convection in the core. You may recall from Chapter 14 that convection is the transport of energy by moving packets of gas. This turbulent motion temporarily keeps more gas from falling in and creates a "surface"—more properly called a photosphere. This photosphere radiates away energy. The hotter it gets, the more energy it radiates, and the bluer that radiation becomes (see Working It Out 5.3). The photosphere of a protostar is tens of thousands of times larger than the surface of the Sun today, and each square meter radiates away energy. As a result, the protostar is thousands of times more luminous than the Sun.

Although the protostar is extremely luminous, astronomers often cannot see it in visible light for two reasons: First, the photosphere of the protostar is relatively cool, so most of its radiation is in the infrared part of the spectrum. Second, and even more important, the protostar is buried deep in the heart of a dense and dusty molecular cloud. However, astronomers are able to view protostars in the infrared part of the spectrum because much of the longer-wavelength infrared light from a protostar is able to escape through the cloud. In addition, as the dust absorbs the visible light, it warms up, and this heated dust also glows in the infrared.

Sensitive infrared instruments developed since the 1980s have revolutionized the study of protostars and other young stellar objects. Dark clouds have revealed themselves to be clusters of dense cloud cores, young stellar objects, and glowing dust when viewed in the infrared. Stars are forming in nodules attached to the tops of the columns of dust and gas in the Eagle Nebula, shown in **Figure 15.16**.

The Evolving Protostar

At any given moment, the protostar is in balance: the forces from hot gas pushing outward and the force of gravity pulling inward exactly oppose each other. However, this balance is constantly changing. Once the core switches from convection to radiation, the deuterium in the core becomes depleted. This allows material to resume falling onto the protostar, adding to its mass and gravitational pull inward and therefore increasing the weight that underlying layers of the protostar must

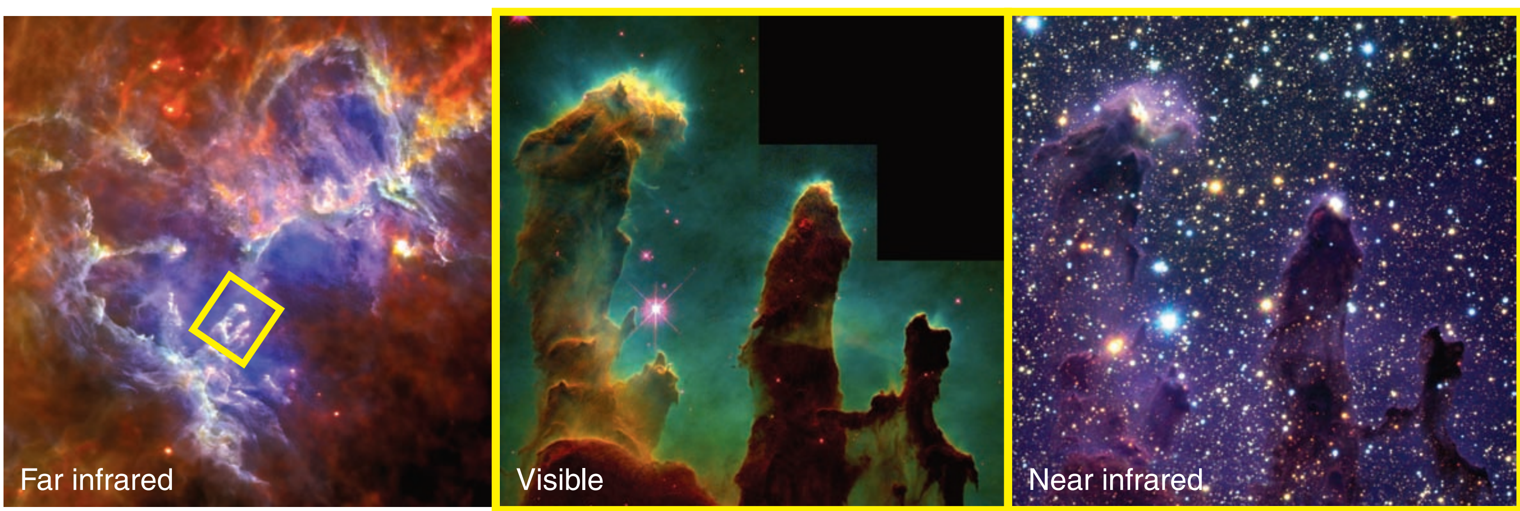


Figure 15.16 The Eagle Nebula contains dense columns of molecular gas and dust at the edge of an H II region. The yellow box in the left image identifies the region that is magnified in the second and third images.

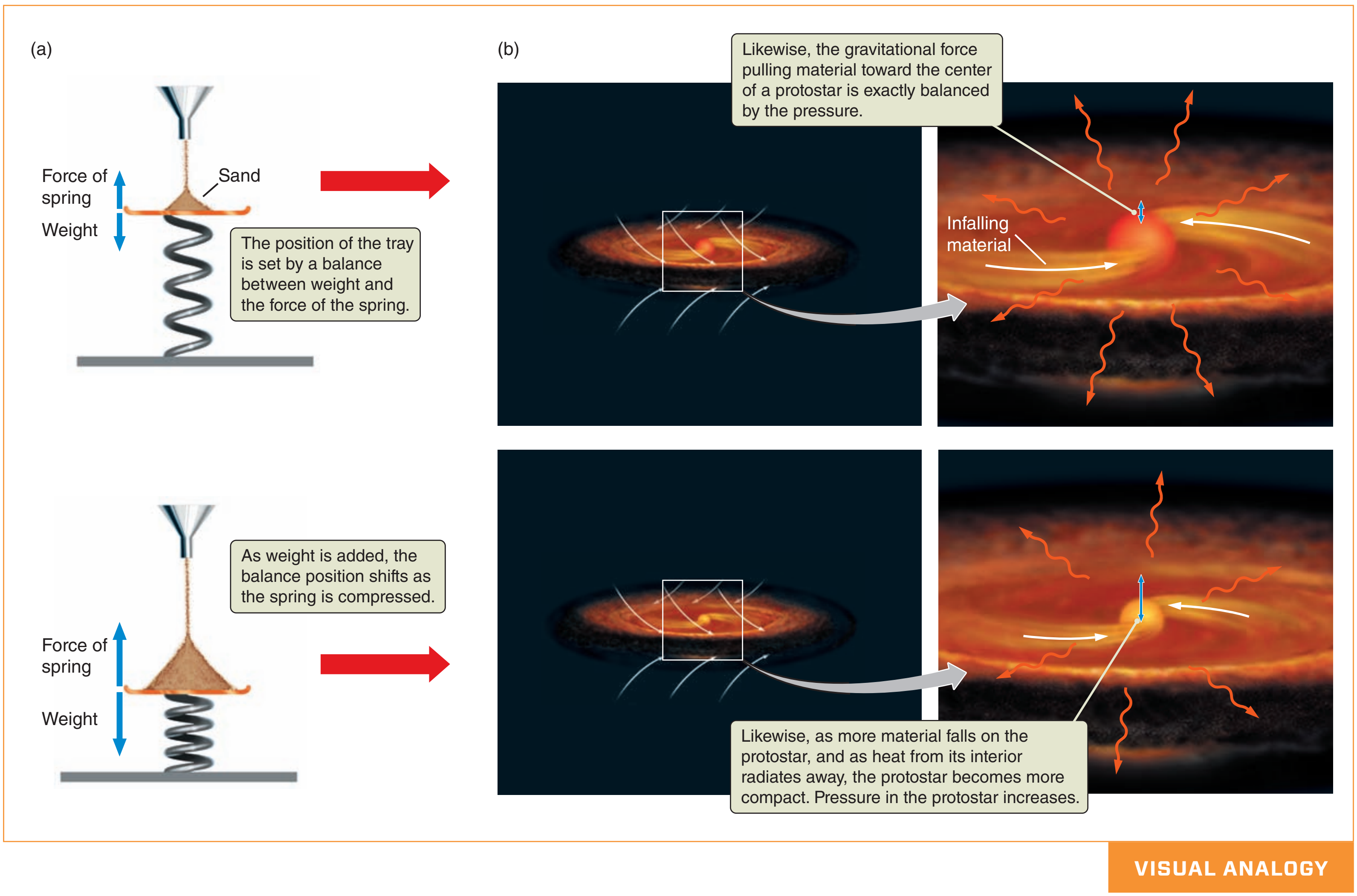
support. The protostar also slowly loses its internal thermal energy by radiating it away.

How can an object be in perfect balance and yet be changing at the same time? Consider an everyday example. **Figure 15.17a** shows a simple spring balance, which works on the principle that the more a spring is compressed, the harder it pushes back. You can measure the weight of an object by determining the point at which the pull of gravity and the push of the spring are equal.

When sand is poured slowly onto the spring balance, at any instant the downward weight of the sand is balanced by the upward force of the spring. As the weight of the sand increases, the spring is slowly compressed. The spring and the weight of the sand are always in balance, but this balance is changing with time as more sand is added. Turning to Figure 15.17b, you can see that the situation is analogous to that of the protostar, in which the outward pressure of the gas behaves like the spring. The inward force of gravity is always matched by the internal pressure pushing out.

Material falls onto the protostar, adding to its mass and gravitational pull. Even though the protostar slowly loses internal thermal energy by radiating it away, the material that has fallen onto the protostar also compresses the protostar and heats it up. The interior becomes denser and hotter, and the pressure rises—just enough

Figure 15.17 (a) A spring balance comes to rest at the point where the downward force of gravity is matched by the upward force of the compressed spring. As sand is added, the location of this balance point shifts. (b) Similarly, the structure of a protostar is determined by a balance between pressure and gravity. Like the spring balance, the structure of the protostar constantly shifts as additional material falls onto its surface and as the protostar radiates energy away.



to balance the increased weight of the material above it. Dynamic balance is always maintained as the protostar slowly contracts.

Figure 15.18 illustrates this chain of events as the protostar shrinks. Gravitational energy is converted to thermal energy, which heats the core, raising the pressure to oppose gravity. This process continues, with the protostar becoming smaller and smaller and its interior growing hotter and hotter. If the protostar is massive enough, its interior will eventually become so hot that nuclear fusion of hydrogen to helium can begin. This is the point at which the transition from protostar to star takes place. The distinction between the two is that a protostar draws its energy from gravitational collapse, whereas a star draws its energy from thermonuclear reactions in its interior.

The protostar's mass determines whether it will actually become a star. As the protostar slowly collapses, the temperature at its center rises. If the protostar's mass is greater than about 0.08 times the mass of the Sun (0.08 M_{Sun}), the temperature in its core will eventually reach 10 million K, and fusion of hydrogen into helium will begin. The newly born star will once again adjust its structure until it is radiating energy away from its surface at just the rate that energy is being liberated in its interior. As it does so, it achieves hydrostatic and thermal equilibrium and "settles" onto the main sequence of the Hertzsprung-Russell (H-R) diagram, where it will spend the majority of its life.

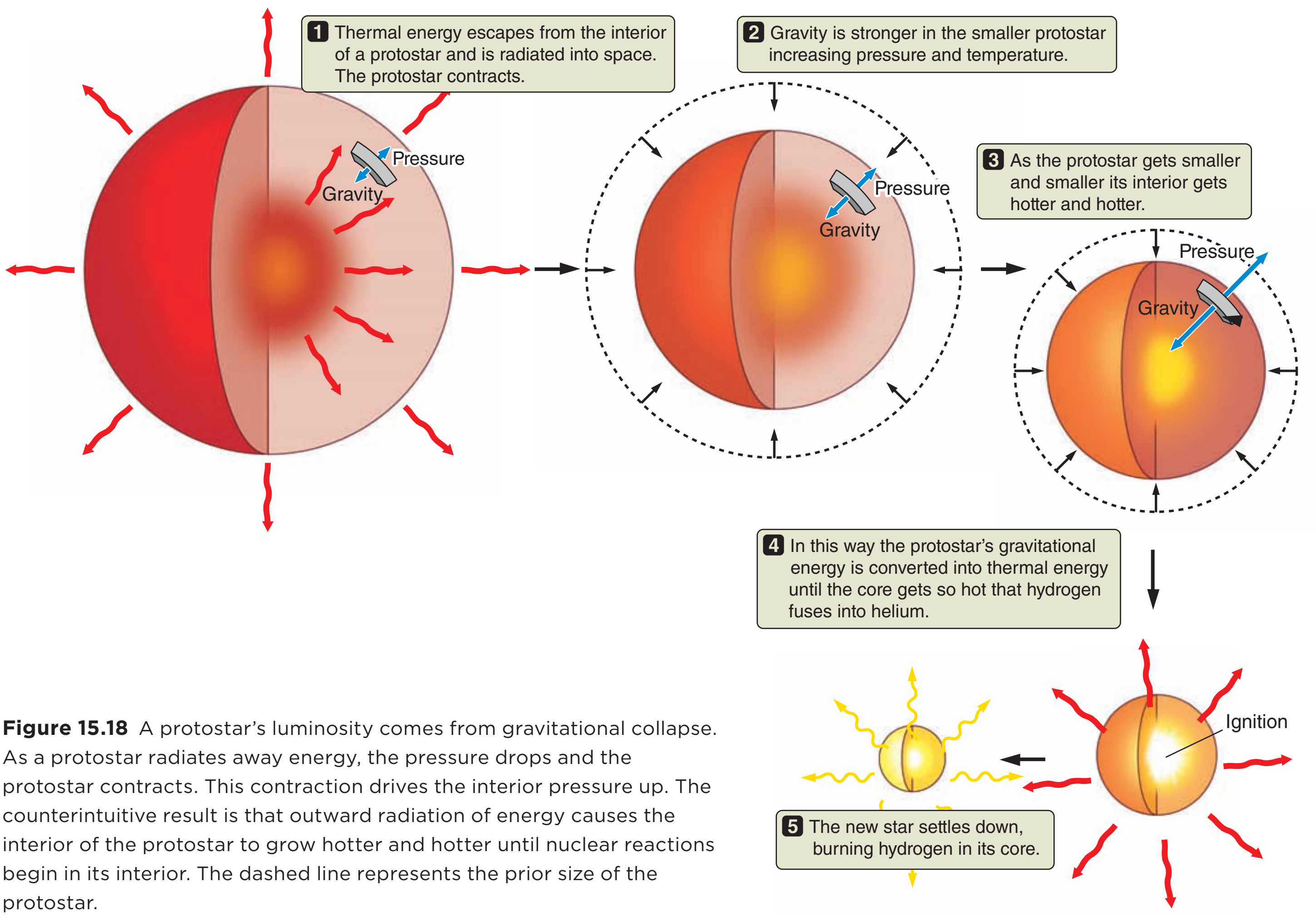


Figure 15.18 A protostar's luminosity comes from gravitational collapse. As a protostar radiates away energy, the pressure drops and the protostar contracts. This contraction drives the interior pressure up. The counterintuitive result is that outward radiation of energy causes the interior of the protostar to grow hotter and hotter until nuclear reactions begin in its interior. The dashed line represents the prior size of the protostar.

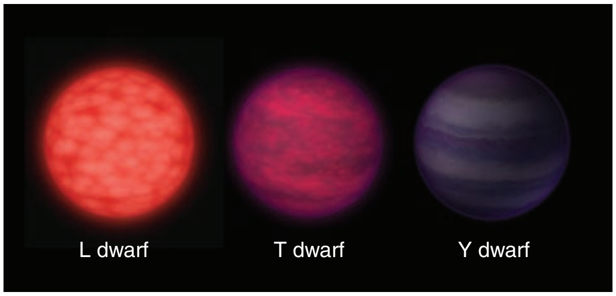


Figure 15.19 This artist's conception shows the three types of brown dwarf stars: L dwarfs (T ~1700 K); T dwarfs (T ~1200 K); and Y dwarfs (T ~500 K).

Brown Dwarfs

If the mass of the protostar is less than 0.08 M_{Sun}, it will never reach the point at which sustained nuclear burning takes place. An object of roughly this mass is called a **brown dwarf**, or sometimes it is called a *substellar object*. A brown dwarf forms in the same way a star forms, yet in many respects it is like a giant planet Jupiter. The International Astronomical Union (IAU) has somewhat arbitrarily set the boundary between a brown dwarf and a supermassive giant planet at 13 Jupiter masses (13 M_{Jup}), although some astronomers think that 10 M_{Jup} would be more appropriate. That is, if the mass of the object is greater than 10 M_{Jup}, it must be a brown dwarf, not a supermassive planet. The upper limit of brown dwarf masses is about 75–80 M_{Jup}. Despite this range of brown dwarf masses, they all have radii about the same as Jupiter's radius—more massive brown dwarfs are denser.

Brown dwarf spectral types L, T, and Y have been added to the sequence of spectral classes to represent stars that are even cooler than M stars (**Figure 15.19**). A brown dwarf never grows hot enough to burn the most common hydrogen nuclei consisting of a single proton, but instead glows primarily by continually cannibalizing its own gravitational energy. The cores of brown dwarfs larger than 13 M_{Jup} can get hot enough to burn deuterium (^{2}H), and those with a mass greater than 65 M_{Jup} can burn lithium. But both of these energy sources are in very short supply, and after a brief period of deuterium or lithium fusion, brown dwarfs shine only by the energy of their own gravitational contraction. As the years pass, a brown dwarf becomes progressively smaller and fainter. The coldest Y dwarfs observed with the WISE infrared space telescope are colder than the human body, which radiates at 310 K.

Since the first brown dwarfs were identified in the mid-1990s, more than a thousand have been found. The cooler among them have methane and ammonia in their atmospheres, similar to what is found in the atmospheres of the giant planets of the Solar System. Winds on brown dwarfs can be very high, producing weather far more violent than storms observed in the atmospheres of the giant planets. Planets have been found orbiting a few brown dwarfs. The nearest one to us is a binary brown dwarf system about 2 parsecs (6.5 light-years) away.

CHECK YOUR UNDERSTANDING 15.3

The energy required to begin nuclear fusion in a protostar originally came from: (a) the gravitational potential energy of the protostar; (b) the kinetic energy of the protostar; (c) the wind from nearby stars; (d) the pressure from the interstellar medium.

15.4 Evolution Before the Main Sequence

Protostars and young evolving stars will change their location on the H-R diagram as they settle into the main sequence, where they will spend the bulk of their lives. In this section, we examine some of the early stages in the life of the new stars.

The Evolutionary Track of an Evolving Star

Within the young protostar, convection carries energy outward, keeping the protostar's interior well stirred. Although the interior of the protostar grows hotter and hotter as it contracts, its photosphere stays at about the same, much cooler, temperature through most of this phase of its evolution. This temperature difference is similar to that between the photosphere of the Sun, which is about 5780 K, and its interior, which is millions of kelvins.

In the 1960s, the theoretical physicist Chushiro Hayashi (1920–2010) explained the difference between the surface temperature of a star or protostar and the temperature deep in its interior. Hayashi showed that the atmospheres of stars and protostars contain a natural thermostat: the H^- ion. (An H^-, or "H minus," ion is a hydrogen atom that has acquired an extra electron and therefore has a negative charge.) The amount of H^- in the atmosphere of a protostar is highly sensitive to the temperature at the protostar's surface. The cooler the atmosphere of a star, the more slowly atoms and electrons are moving, and the easier it is for a hydrogen atom to hold on to an extra electron. As a result, the cooler the atmosphere of the star, the more H^- there is.

The H^- ion, in turn, helps control how much energy a star or protostar radiates away. The more H^- there is in the atmosphere of the star or protostar, the more opaque the atmosphere is, and the more effectively the thermal energy of the protostar is trapped in its interior. Imagine that the surface of the protostar is "too cool," meaning that extra H^- forms in the atmosphere and makes the atmosphere of the protostar more opaque. The atmosphere thus traps more of the radiation that is trying to escape, and the trapped energy heats up the star. As the temperature climbs, the H^- ions are changed to neutral H atoms. Now imagine the other possibility—that the protostar is too hot. In this case, H^- in the protostar's atmosphere is destroyed, so the atmosphere becomes more transparent, allowing radiation to escape more freely from the interior. Because the protostar cannot hold on to enough of its energy to stay warm, the surface cools. In either case—too cold or too hot—H^- is formed or destroyed until the star's atmosphere once again traps just the right amount of escaping radiation. The H^- ion is basically doing the same thing that you do with your bedcovers at night. If you get too cold, you pile on extra covers to trap your body's thermal energy and keep warm (more H^- ions). If you get too hot, you kick off some covers to cool off (fewer H^- ions).

The amount of H^- in the atmosphere keeps the surface temperature of the protostar somewhere between about 3000 and 5000 K, depending on the protostar's mass and age. Because the surface temperature of the protostar is not changing much, the amount of energy per unit time (power) radiated away by each square meter of the surface of the protostar does not change much either. Recall the Stefan-Boltzmann law from Chapter 5, which says that the amount radiated by each square meter of an object's surface is determined by its temperature. As the protostar shrinks, the area of its surface shrinks as well. There are fewer square meters of surface from which to radiate, so the luminosity of the protostar drops. As viewed from the outside, the protostar stays at nearly the same temperature and color but gradually gets fainter as it evolves toward its eventual life as a main-sequence star.

In Chapter 13, we introduced the H-R diagram and used it to help explain how the properties of stars differ. For the next several chapters, we will use the H-R diagram to keep track of how stars change as they evolve through their lifetimes. The path on the H-R diagram that a star follows as it goes through the different stages of its life is called the star's **evolutionary track**. The protostar is brighter than it will be as a true star on the main sequence, so a protostar's track is located above the main sequence on the H-R diagram. **Figure 15.20** shows the pre-main-sequence evolutionary tracks of stars of several different

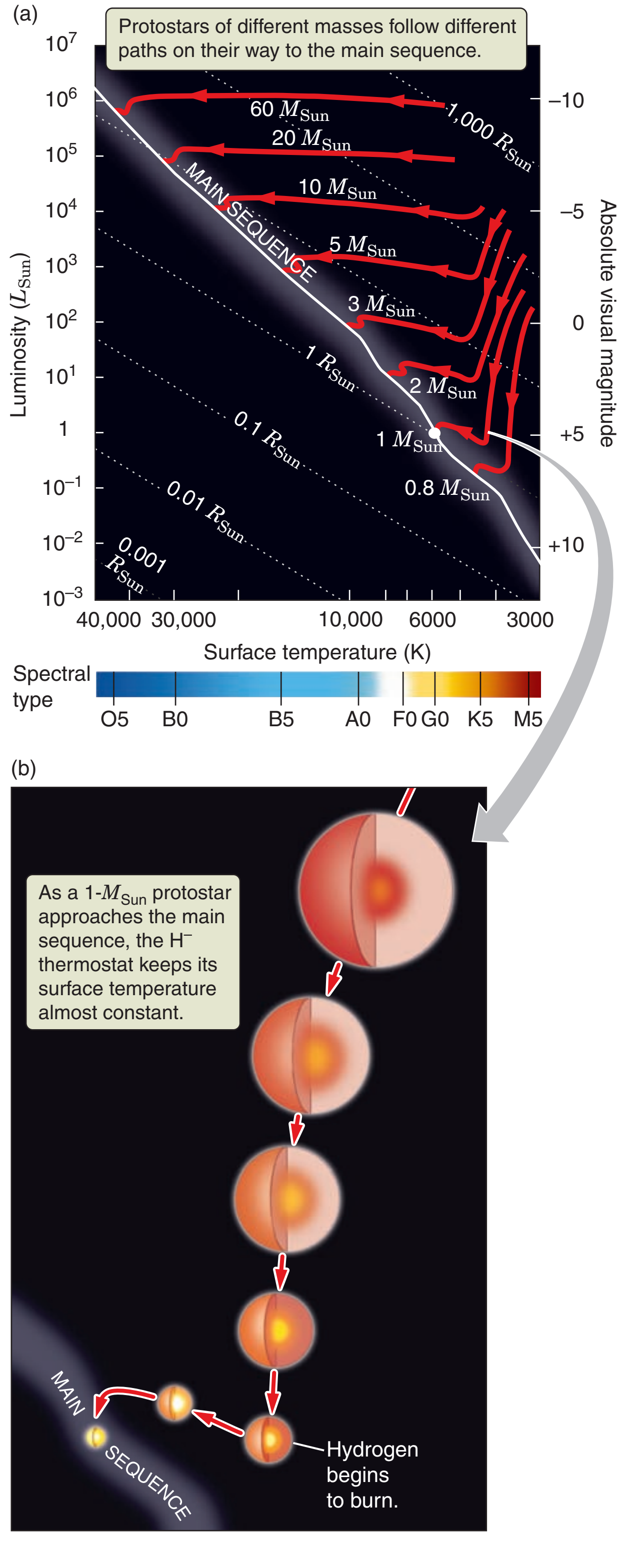


Figure 15.20 (a) The evolution of pre-main-sequence stars can be followed on the H-R diagram. More massive protostars in the upper right portion of the diagram follow horizontal tracks. (b) The roughly vertical, constant-temperature part of the evolutionary track of a low-mass protostar is called the Hayashi track.

15.2 Working It Out Luminosity, Surface Temperature, and Radius of Protostars

In Chapter 13, you learned how the luminosity, surface temperature, and radius of a star are related:

$$L = 4\pi R^2 \sigma T^4$$

What can this equation reveal about the changing properties of the protostar as it shrinks its radius?

Suppose that when the Sun was a protostar, it had a radius 10 times what it is now and a surface temperature of 3300 K. What would its luminosity have been? The equations for each are

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

and

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

We can set this up as a ratio, comparing the luminosity of the protostar Sun with its luminosity now, L_{Sun}:

$$\frac{L_{\text{protostar}}}{L_{\text{Sun}}} = \frac{4\pi R^2_{\text{protostar}} \sigma T^4_{\text{protostar}}}{4\pi R^2_{\text{Sun}} \sigma T^4_{\text{Sun}}}$$

We rewrite this as follows, grouping like terms together:

$$\frac{L_{\text{protostar}}}{L_{\text{Sun}}} = \frac{4\pi\sigma}{4\pi\sigma} \times \left(\frac{R_{\text{protostar}}}{R_{\text{Sun}}}\right)^2 \times \left(\frac{T_{\text{protostar}}}{T_{\text{Sun}}}\right)^4$$

Then we cancel out the constants, $4\pi\sigma$, and use the value for $T_{\text{Sun}} = 5780$ K from Chapter 14. We know that the protostar's radius is 10 times that of the Sun, so $R_{\text{protostar}}/R_{\text{Sun}} = 10$. Then the equation becomes

$$\frac{L_{\text{protostar}}}{L_{\text{Sun}}} = \left(\frac{10}{1}\right)^2 \times \left(\frac{3300}{5780}\right)^4 = 10^2 \times (0.57)^4 = 10.6$$

So the Sun was about 10.6 times more luminous as a protostar than it is now. We see this on the H-R diagram of protostars (see Figure 15.20). As a 1-M_{Sun} star approaches the main sequence on the diagram, it moves down (toward lower luminosity) and to the left (toward higher surface temperature).

masses. In a higher-mass protostar, the luminosity stays about the same, but the surface temperature increases; on the H-R diagram, its track moves left horizontally to the main sequence. In a lower-mass protostar, the temperature stays about the same, but the luminosity decreases. On the H-R diagram, its track moves vertically down to the main sequence; this is called its **Hayashi track** (Figure 15.20b). The relationship between surface temperature, luminosity, and radius of protostars is further explored in **Working It Out 15.2**.

Bipolar Outflow

As shown in **Figure 15.21a**, material falls onto the accretion disk around a young stellar object and moves inward toward the equator of the star, while at the same time other material is blown away from the protostar and disk in two opposite directions from the plane of the disk. The resulting stream of material away from the protostar is called a **bipolar outflow**. Powerful outflows can disrupt the cloud core and accretion disk from which the protostar formed, shutting down the flow of material onto the protostar.

Some bipolar outflows from young stellar objects are slow and fairly disordered, but others produce remarkable **jets** of material that move away from the central protostar and disk at velocities of hundreds of kilometers per second (Figure 15.21b). The material in these jets flows out into the interstellar medium, where it heats, compresses, and pushes away surrounding interstellar gas. Knots of glowing gas accelerated by jets are referred to as **Herbig-Haro objects** (or **HH**

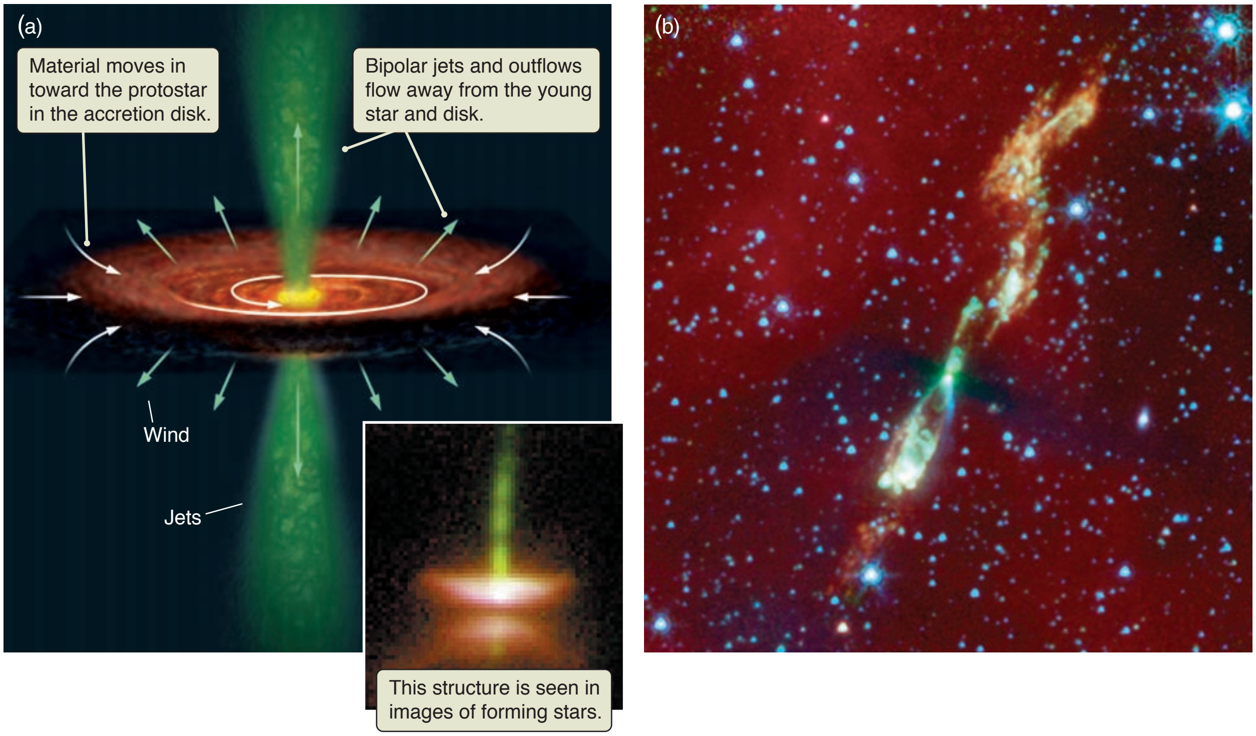


Figure 15.21 (a) Material falls onto an accretion disk around a protostar and then moves inward, eventually falling onto the protostar. In the process, some of this material is driven away in powerful jets that stream perpendicular to the disk. (b) This infrared Spitzer Space Telescope image shows jets streaming outward from a young, developing star. Note the nearly edge-on, dark accretion disk surrounding the young star.

objects for short), named after the two astronomers who first identified them and associated them with star formation. An example is shown in **Figure 15.22**.

The origin of outflows from protostars is not well understood, but current models suggest that they are the result of magnetic interactions between the protostar and the disk. The interior of a protostar on its Hayashi track is convective. Great cells of hot gas are rising from the interior, while other cells of cooler gas are falling toward the center. This convection, coupled with the protostar's rapid rotation, can lead to the formation of a dynamo, similar to the dynamo that drives the Sun's magnetic field. The dynamo in the center of a protostar would be much more powerful than the Sun's dynamo, however. The protostar's resulting strong magnetic field might cause the protostar to begin blowing a powerful wind. It might also act something like the blade in a blender, tearing at the inner edge of the accretion disk and flinging material off into interstellar space.

Until the protostellar wind begins, the protostar is enshrouded in the dusty molecular-cloud core from which it was born. As the wind from the protostar disperses this obscuring envelope, the first direct, visible-light view of the protostar emerges: the protostar is "revealed." Some of these protostars of lower mass are called **T Tauri stars**. This name comes from the first recognized member of this class of objects, the star labeled T in the constellation Taurus. Higher-mass

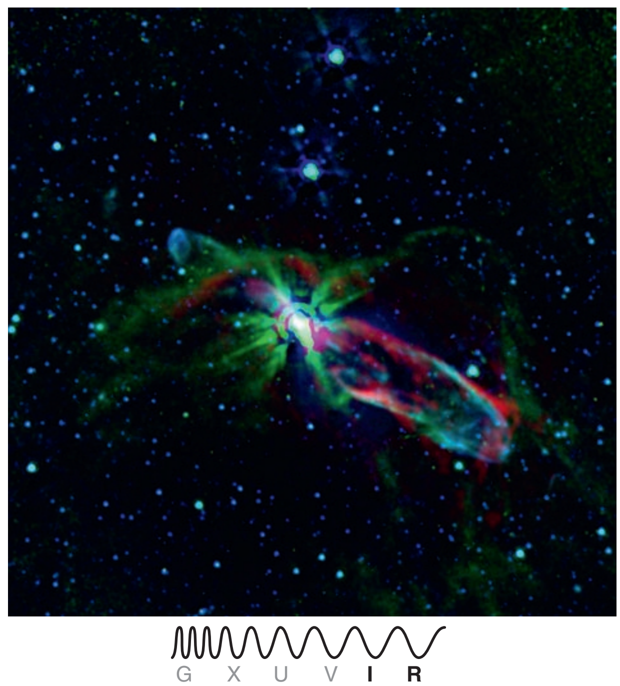


Figure 15.22 In this combined Spitzer infrared and Atacama Large Millimeter/submillimeter Array (ALMA) observation of Herbig-Haro object HH 46/47, twin supersonic jets originate from the newborn central star, blasting away surrounding gas and forming two lobes.

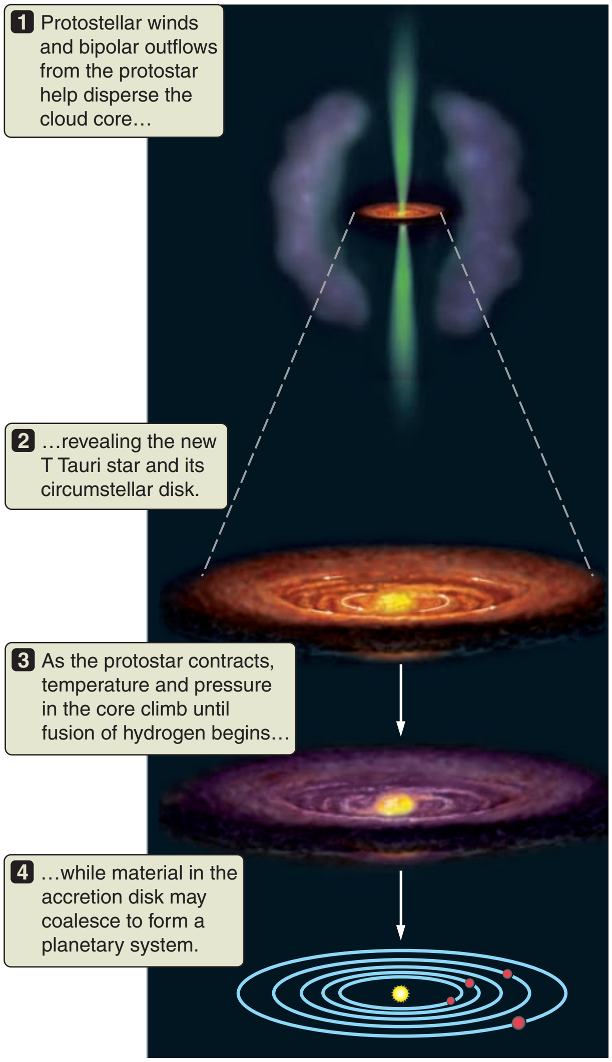


Figure 15.23 This figure presents an overview of how stars like the Sun form, beginning with the onset of stellar winds and ending with the ignition of a star sitting at the center of a revolving system of planets.

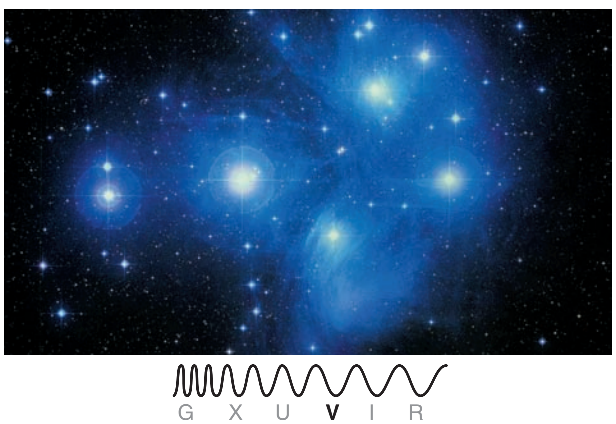


Figure 15.24 The Pleiades or "Seven Sisters" is a cluster of young stars. The diffuse blue light around the stars is starlight scattered by interstellar dust.

protostars of spectral type B or A are called Herbig Be/Ae stars. **Figure 15.23** summarizes the evolution of a protostar from the onset of winds to the formation of planets.

The Influence of Mass

Astronomers are very interested in how and why molecular clouds subdivide themselves into stars with a range of masses. The details of this division—specifically, what fraction of newly formed stars will have which masses—are crucial if observations of the stars in the Sun's vicinity today are to help untangle the history of star formation in our galaxy. Astronomers do not understand why some cloud cores become 1-M_{Sun} stars while others become 5- or 10-M_{Sun} stars.

A look around our galaxy at large reveals a variety of stars—some very old and others very young. If these were the only stars available to study, it would be extremely difficult to learn much about how stars evolve. But astronomers have long known that stars are often found together in close collections called **star clusters**. **Figure 15.24** shows one such star cluster: a group called the Pleiades or "Seven Sisters." Star clusters are collections of stars that all formed in the same place, from the same material, and at about the same time. They provide extremely useful samples for studying star formation. Even though the few brightest and most massive stars in a cluster dominate any observation of a cluster, most of the stars in a cluster are less massive than the Sun. In fact, some star-forming regions do not seem to form any especially massive stars at all.

After a cloud-core collapse, the evolution of a protostar is determined largely by its mass. Calculations suggest that a star like the Sun probably takes about 10 million years or so to descend its Hayashi track and become a star on the main sequence. Taking into account the entire history of protostellar formation, including the collapse and fragmentation of the molecular cloud itself, the total time for initial fragmentation into cloud core up to the ignition of hydrogen burning might be more like 30 million years. Because the self-gravity of a more massive core is stronger, more massive cores collapse to form stars more quickly. A 10-M_{Sun} star might go from the stage of being a molecular-cloud core to that of burning hydrogen in its interior in only 100,000 years. A 100-M_{Sun} star might take less than 10,000 years. By comparison, a 0.1-M_{Sun} star might take 100 million years finally to reach the main sequence.

The 30 million years or so that it took for the Sun to form is a long time, but it is a tiny fraction of the 10 billion years during which the Sun will steadily fuse hydrogen into helium as a main-sequence star. It is no wonder that so few among the many stars visible in the sky are young. But every star was young at one time, including the Sun.

CHECK YOUR UNDERSTANDING 15.4

Why are so few of the many stars that astronomers see in the sky protostars?
(a) Protostars are hidden in giant molecular clouds. (b) Protostars are small.
(c) Protostars are dim. (d) Protostars are short-lived.

Origins

Star Formation, Planets, and Life

When astronomers consider the possibility of other life in the universe, one of the first things they think about is the formation of stars and planets. Life probably needs planets, and planets form along with stars. The conditions under which a star is born, and the mass and chemical composition that it has when it begins its nuclear fusion, set the stage for the rest of its life. In Chapter 7, we noted that planet formation can be a common consequence of star formation; thus, if a star is going to have planets, they form at about the same time as the star. If the star is going to have rocky planets with hard surfaces (such as the planets of the inner Solar System) or gaseous planets with rocky cores and rocky moons (such as the planets of the outer Solar System), then the material from which the star and planets form must be "enriched" with the heavy elements that make up these rocky surfaces.

These enriched clouds would also provide elements that are essential to life on Earth. In addition to the presence of organic molecules mentioned earlier in the chapter, astronomers have detected water in star-forming regions such as W3 IRS5 (**Figure 15.25**). Water exists as ice mixed with dust grains in the cool molecular clouds or as vapor when it is closer to a protostar and the dust grains and ice evaporate. In 2011, the Herschel Space Observatory detected oxygen molecules (O_2, the type we breathe) in a star-forming complex in Orion. Oxygen is the third most common element in the universe, yet it had not been decisively observed before in molecular form. This oxygen also may have come from the melting and evaporation of water ice on the tiny dust grains.

As noted in Chapter 13, astronomers had doubted that planets could exist in stable orbits in binary star systems, but now a few such circumbinary systems have been found. Planets that form within associations of O and B stars may be too unstable to last very long. Isolated planets unattached to any star likely move through the Milky Way. Perhaps these isolated planets were gravitationally ejected soon after their formation in a multiple system. But these planets do not have a source of energy like Earth's Sun. Astronomers theorize that only planets that orbit stars will be able to support life. So when they try to estimate the possibility of life in the galaxy, astronomers include estimations of the rate of formation of stars in the galaxy, along with the fraction of stars that have planets. Advances in the study of star formation and planet detection help astronomers to understand better the conditions under which life might develop elsewhere.

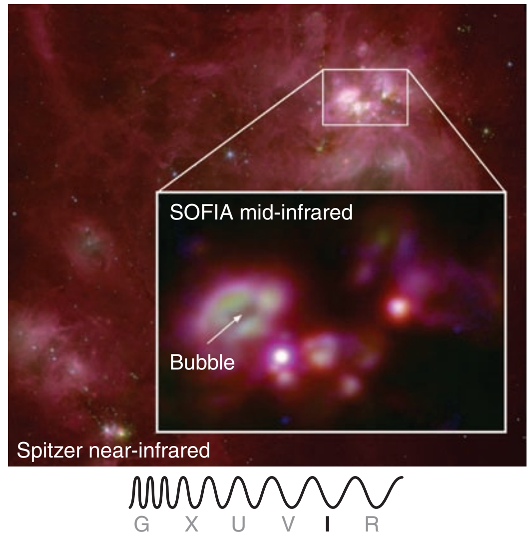


Figure 15.25 Water was detected in the W3 IRS5 star-forming complex. Here, W3 IRS5 is observed with the Spitzer Space Telescope in near infrared and with the Stratospheric Observatory for Infrared Astronomy (SOFIA) telescope in mid-infrared. A massive star has cleared the dust and gas from a small bubble, sweeping it into a dense shell (green).

ARTICLES QUESTIONS

In this article, astronomers report on their search for dust from outside of the Solar System.

Interstellar Dust Discovered Inside NASA Spacecraft

By **IRENE KLOTZ, Discovery News**

Thanks to a massive effort by 30,716 volunteers, scientists have pinpointed what appear to be seven precious specks of dust from outside the Solar System, each bearing unique stories of exploded stars, cold interstellar clouds, and other past cosmic lives.

The Herculean effort began eight years ago after NASA's *Stardust* robotic probe flew by Earth to deposit a capsule containing samples from a comet and dust grains from what scientists hoped would be interstellar space. The spacecraft was outfitted with panels containing a smoke-like substance called aerogel that could trap and preserve fast-moving particles.

Stardust twice put itself into position to fish for interstellar grains, which are so small that a trillion of them would fit in a teaspoon. The only way scientists back on Earth would be able to find them was by the microscopic trails the grains made as they plowed into the aerogel.

"When we did the math we realized it would take us decades to do the search ourselves," physicist Andrew Westphal, with the University of California, Berkeley, told Discovery News.

The team used an automated microscope to scan the collector and put out a call for volunteers.

"This whole approach was treated with pretty justifiable criticism by people in my community. They said, 'How can you trust total strangers to take on this project?'" Westphal said.

"We really didn't know how else to do it. We still don't," he added.

Recruits were trained and had to pass a test before they were given digital scans to peruse. Scientists sometimes inserted images with known trails just to see if the volunteers, known as "dusters," would spot them.

"We were very pleased to see that people are really good at finding these tracks, even really, really difficult things to find," Westphal said.

More than 50 candidate dust motes turned out to be bits of the spacecraft itself, but scientists found seven specks that bear chemical signs of interstellar origin and travel.

The grains are surprisingly diverse in shape, size, and chemical composition. The larger ones, for example, have a fluffy, snowflake-like structure.

Additional tests are needed to verify the grains' interstellar origins and ferret out their histories. But the grains are so tiny that with currently available technology, additional analysis would mean their demise.

"It'll probably be years before we can do a lot more with these samples," said space scientist Mike Zolensky, who oversees NASA's collection of cosmic dust, moon rocks, and other extraterrestrial samples at the Johnson Space Center in Houston.

"But we've got them safely tucked away and we can hang on to them until those techniques come along," Zolensky said.

ARTICLES QUESTIONS

1. Why do scientists want to identify interstellar dust grains?
2. How were these dust grains distinguished from Solar System grains?
3. Why can't the scientists do a complete analysis on these particles?
4. How did volunteer "citizen scientists" assist with this project?
5. Visit the website noted in question 49 in the "Questions and Problems" section later in this chapter. Is this project continuing? What else has been discovered?

"Interstellar dust discovered inside NASA spacecraft," by Irene Klotz. Discovery News, August 14, 2014. Reprinted by permission of Discovery Access.

Summary

Astronomers have never watched the full process of a star forming from beginning to end. Instead, they have observed many different stars at different stages in their formation and evolution at different wavelengths, and they have used their knowledge of physical laws to tie these observations together into a coherent, consistent description of how, why, and where stars form. Stars form from material in the interstellar medium, under the influence of gravity, often triggered by external factors such as the winds of hot stars nearby. The conditions under which a star is born determine whether it will have planets along with the chemical elements required by life as it exists on Earth.

LG 1 **Describe the types and states of material that exist in the space between the stars and how this material is detected.** The interstellar medium is complex, ranging from cold, relatively dense molecular clouds to hot, tenuous intercloud gas heated and ionized by energy from stars and stellar explosions. Dust and gas in the interstellar medium blocks much visible light but becomes more transparent at longer, infrared wavelengths. Different phases of the interstellar medium emit various types of radiation and can be observed at different wavelengths, ranging from radio waves to X-rays. Neutral hydrogen cannot be detected at visible and infrared wavelengths, but it is revealed by its 21-cm emission.

LG 2 **Explain the conditions under which a cloud of gas can contract into a stellar system and the role that gravity and angular momentum play in the formation of stars and planets.** Star formation begins when the self-gravity of dense clouds exceeds outward pressure. The clouds collapse, heat up, and fragment to form stars. The conservation of angular momentum is important to the formation of disks during the collapse. Forming stars are detected from their infrared emission and from the effects they have on their surroundings.

LG 3 **List the steps in the evolution of a protostar.** Protostars collapse, radiating away their gravitational energy until fusion starts in their cores. When they reach hydrostatic and thermal equilibrium, they settle onto the main sequence. Stars form in clusters from dense cores buried within giant molecular clouds. A protostar must have a mass of at least 0.08 M_{Sun} to become a true star. Brown dwarfs are neither stars nor planets, but something in between.

LG 4 **Describe the track of a protostar as it evolves to a main-sequence star on the Hertzsprung-Russell (HR) diagram.** Because star formation takes tens of thousands to millions of years, what astronomers know about the evolution of the birth of stars comes from observations of many protostars at various stages of their development. Protostars of different masses follow different paths to the main sequence, but in general there is a stage at which the protostar moves horizontally, at roughly constant luminosity, followed by a constant temperature drop nearly straight down toward the main sequence. Once hydrogen burning begins, the star again moves nearly horizontally until it reaches the main sequence.

? UNANSWERED QUESTIONS

- Many questions about star formation remain. For example, how must theories be modified to explain the formation of binary stars or other multiple-star systems? At what point during star formation is it determined that a collapsing cloud core will form several stars instead of just one? Some models suggest that this split may happen early in the process, during the fragmentation and collapse of the molecular cloud. The advantage of these ideas is that they provide a natural way of dealing with much of the angular momentum of the cloud core: it goes into the orbital angular momentum of the stars around each other. Other models suggest that additional stars may form from the accretion disk around an initially single protostar.
- Do high-mass and low-mass stars form in very different ways? The smallest stars, spectral type M, are most likely to form as single stars, but a high fraction of medium-mass stars are formed in binary pairs. One theory is that these binaries actually start out as triple systems, from which the smallest star is gravitationally ejected, leading to a remaining pair and a single star. The highest-mass stars are less likely to form alone; many form in OB associations—larger groups of massive stars in which the formation of one large star may stimulate the formation of another nearby in the molecular cloud.
- How common are brown dwarfs? As with extrasolar planets, they have been observed only recently, so their space density is not well known. The Kepler Mission, which finds extrasolar planets (see Chapter 7) and eclipsing binaries (see Chapter 13), also detects brown dwarfs, so there may be an answer to this question in a few years.

Questions and Problems

Test Your Understanding

1. Phases of the interstellar medium include (choose all that apply)
   a. hot, low-density gas.
   b. cold, high-density gas.
   c. hot, high-density gas.
   d. cold, low-density gas.

2. Dust in the interstellar medium can be observed in
   a. visible light.
   b. infrared radiation.
   c. radio waves.
   d. X-rays.

3. The interstellar medium in the Sun's region of the galaxy is closest in composition to
   a. the Sun.
   b. Jupiter.
   c. Earth.
   d. comets in the Oort Cloud.

4. Interstellar dust is effective at blocking visible light because
   a. the dust is so dense.
   b. dust grains are so few.
   c. dust grains are so small.
   d. dust grains are so large.

5. Hot intercloud gas is heated primarily by
   a. starlight.
   b. protostars.
   c. supernova explosions.
   d. neutrinos.

6. Astronomers determined the composition of the interstellar medium from
   a. observing its emission and absorption lines.
   b. measuring the composition of the planets.
   c. return samples from spacecraft.
   d. composition of meteorites.

7. In astronomy, the term *bipolar* refers to outflows that
   a. point in opposite directions.
   b. alternate between expanding and collapsing.
   c. rotate about a polar axis.
   d. show spiral structure.

8. Which of the following has contributed most to our understanding of the process of star formation?
   a. Astronomers have observed star formation as it happens for a small number of stars.
   b. Astronomers have observed star formation as it happens for a large number of stars.
   c. Astronomers have observed many different stars at each step of the formation process.
   d. Theoretical models predict the way stars form.

9. Cold neutral hydrogen can be detected because
   a. it emits light when electrons drop through energy levels.
   b. it blocks the light from more distant stars.
   c. it is always hot enough to glow in the radio and infrared wavelengths.
   d. the atoms in the gas change spin states.

10. The Hayashi track is a nearly vertical evolutionary track on the H-R diagram for low-mass protostars. Which of the following would you expect from a protostar moving along a vertical track?
    a. The star remains the same brightness.
    b. The star remains the same luminosity.
    c. The star remains the same color.
    d. The star remains the same size.

11. Which two forces establish hydrostatic equilibrium in an evolving protostar?
    a. the force from pressure and gravity
    b. the force from pressure and the strong nuclear force
    c. gravity and the strong nuclear force
    d. energy emitted and energy produced

12. Suppose you are studying a visible-light image of a distant galaxy, and you see a dark lane cutting across the bright disk. This dark line is most likely caused by
    a. gravitational instabilities that clear the area of stars.
    b. dust in the Milky Way blocking the view of the distant galaxy.
    c. dust in the distant galaxy blocking the view of stars in the disk.
    d. a flaw in the instrumentation.

13. What causes a hydrogen atom to radiate a photon of 21-cm radio emission?
    a. The electron drops down one energy level.
    b. The formerly free electron is captured by the proton.
    c. The electron flips to an aligned spin state.
    d. The electron flips to an unaligned spin state.

14. Astronomers know that there are dusty accretion disks around protostars because
    a. there is often a dark band across the protostar.
    b. there is often a bright band across the protostar.
    c. theory says accretion disks should be there.
    d. there are planets in the Solar System.

15. What is the single most important property of a star that will determine its evolution?
    a. temperature
    b. composition
    c. mass
    d. radius

Thinking about the Concepts

16. The interstellar medium is approximately 99 percent gas and 1 percent dust. Why does dust and not gas block a visible-light view of the galactic center?

17. Explain why observations in the infrared are necessary for astronomers to study the detailed processes of star formation.

18. How does the material in interstellar clouds and intercloud gas differ in density and distribution?

19. When a star forms inside a molecular cloud, what happens to the cloud? Is it possible for a molecular cloud to remain cold and dark with one or more stars inside it? Explain your answer.

20. If you placed your hand in boiling water (100°C) for even 1 second, you would get a very serious burn. If you placed your hand in a hot oven (200°C) for a second or two, you would hardly feel the heat. Explain this difference and how it relates to million-kelvin regions of the interstellar medium.

21. How do astronomers know that the Sun is located in a "local bubble" formed by a supernova?

22. Interstellar gas atoms typically cool by colliding with other gas atoms or grains of dust; during the collision, each gas atom loses energy and hence its temperature is lowered. How does this explain why very low-density gases are generally so hot, while dense gases tend to be so cold?

23. Explain how the 21-cm line discussed in the Process of Science Figure supports the cosmological principle (which states that the laws of physics must be the same everywhere).

24. Molecular hydrogen is very difficult to detect from the ground, but astronomers can easily detect carbon monoxide (CO) by observing its 2.6-cm microwave emission. Describe how observations of CO might help astronomers infer the amounts and distribution of molecular hydrogen within giant molecular clouds.

25. The Milky Way contains several thousand giant molecular clouds. Describe a giant molecular cloud and its role in star formation.

26. As a cloud collapses to form a protostar, the forces of gravity felt by all parts of the cloud (which follow an inverse square law) become stronger and stronger. One might argue that under these conditions, the cloud should keep collapsing until it becomes a single massive object. Why doesn't this happen?

27. The internal structure of a protostar maintains hydrostatic equilibrium even as more material is falling onto it. Explain how this can be.

28. What are the similarities and differences between a brown dwarf and a giant planet such as Jupiter? Would you classify a brown dwarf as a supergiant planet? Explain your answer.

29. The H^- ion acts as a thermostat in controlling the surface temperature of a protostar. Explain the process.

30. How does the composition of the molecular cloud affect the type of planets and stars that form within it?

Applying the Concepts

31. In Chapter 13, you learned that astronomers can measure the temperature of a star by comparing its brightness in blue and yellow light. Does reddening by interstellar dust affect a star's temperature measurement? If so, how?

32. When a hydrogen atom is ionized, it splits into two components.
    a. Identify the two components.
    b. If both components have the same kinetic energy, which moves faster?

33. Estimate the typical density of dust grains (grains per cubic centimeter) in the interstellar medium. A typical grain has a mass of about 10^{-17} kilogram (kg). (Hint: You know the typical density of gas, and the fraction of the interstellar medium's mass that is made of dust.)

34. Referring to Figure 15.3, estimate the blackbody temperature of the star as shown in part (b) (without dust) and part (c) (with dust). How significant are the effects of interstellar dust when observed data are used to determine the properties of a star?

35. A typical temperature of intercloud gas is 8000 K. Using Wien's law (see Working It Out 15.1 and Chapter 5), calculate the wavelength at which this gas would radiate.

36. Some parts of the Orion Nebula have a blackbody peak wavelength of 0.29 μm. What is the temperature of these parts of the nebula?

37. Stellar radiation can convert atomic hydrogen (H I) to ionized hydrogen (H II).
    a. Why does a B8 main-sequence star ionize far more interstellar hydrogen in its vicinity than does a K0 giant of the same luminosity?
    b. What properties of a star are important in determining whether it can ionize large amounts of nearby interstellar hydrogen?

38. The mass of a proton is 1,850 times the mass of an electron. If a proton and an electron have the same kinetic energy ($E_K = \frac{1}{2}mv^2$), how many times greater is the velocity of the electron than that of the proton?

39. If a typical hydrogen atom in a collapsing molecular-cloud core starts at a distance of 1.5×10^{12} km (10,000 AU) from the core's center and falls inward at an average velocity of 1.5 km/s, how many years does it take to reach the newly forming protostar? Assume that a year is 3×10^7 seconds.

40. The ratio of hydrogen atoms (H) to carbon atoms (C) in the Sun's atmosphere is approximately 2,400:1 (see Table 13.1). It would be reasonable to assume that this ratio also applies to molecular clouds. If 2.6-cm radio observations indicate 100 M_{Sun} of carbon monoxide (CO) in a giant molecular cloud, what is the implied mass of molecular hydrogen (H_2) in the cloud? (Carbon represents $\frac{3}{7}$ of the mass of a CO molecule.)

41. Neutral hydrogen emits radiation at a radio wavelength of 21 cm when an atom drops from a higher-energy spin state to a lower-energy spin state. On average, each atom remains in the higher energy state for 11 million years (3.5×10^{14} seconds).
    a. What is the probability that any given atom will make the transition in 1 second?
    b. If there are 6×10^{59} atoms of neutral hydrogen in a 500-M_{Sun} cloud, how many photons of 21-cm radiation will the cloud emit each second?
    c. How does this number compare with the 1.8×10^{45} photons emitted each second by a solar-type star?

42. The Sun took 30 million years to evolve from a collapsing cloud core to a star, with 10 million of those years spent on its Hayashi track. It will spend a total of 10 billion years on the main sequence. Suppose the Sun's main-sequence lifetime were compressed into a single day.
    a. How long would the total collapse phase last?
    b. How long would the Sun spend on its Hayashi track?

43. A protostar with the mass of the Sun starts out with a temperature of about 3500 K and a luminosity about 200 times larger than the Sun's current value. Estimate this protostar's size and compare it to the size of the Sun today.

44. The star-forming region 30 Doradus is 160,000 light-years away in the nearby galaxy called the Large Magellanic Cloud and appears about one-sixth as bright as the faintest stars visible to the naked eye. If it were located at the distance of the Orion Nebula (1,300 light-years away), how much brighter than the faintest visible stars would it appear?

45. Assume a brown dwarf has a surface temperature of 1000 K and approximately the same radius as Jupiter. What is its luminosity compared to that of the Sun? How many brown dwarfs like this one would be needed to produce the luminosity of a star like the Sun?

USING THE WEB

46. Go to the Astronomy Picture of the Day (APOD) website (http://apod.nasa.gov/apod), do a search on "molecular clouds," and pick out a few images. Were these pictures obtained from space or on the ground, and at what wavelengths? With which telescopes? What wavelengths do the colors in the images represent? Are they "real" or "false-color" images?

47. Go to NASA's Spitzer Space Telescope website (http://www.spitzer.caltech.edu). Click on "News" and find a recent story about star formation. What did Spitzer observe? What wavelengths do the colors in the picture represent? How does this "false color" help astronomers to analyze these images? Why do astronomers study star formation in the infrared rather than in the visual part of the spectrum?

48. The Stratospheric Observatory for Infrared Astronomy (SOFIA) is a 2.5-meter telescope on a modified Boeing 747 aircraft. Go to the SOFIA website (http://sofia.usra.edu). Why have astronomers put an infrared telescope on an airplane? What has been detected with this telescope?

49. Citizen science: Go to the website for *Stardust* (http://stardustathome.ssl.berkeley.edu), a project in which volunteers use a virtual microscope to analyze digital scans of particles collected by the *Stardust* mission in 2006. The goal is to identify tiny interstellar dust grains. Follow the steps under "Get Started" (you need to create a log-in account) and help search for stardust. Click on "News." What has been learned from this project? Remember to save the images for your homework, if required.

50. Do a news search for a story about brown dwarfs. Is this story from an observatory? A NASA mission? A press release? What is new, and why is it interesting?

smartwork5

If your instructor assigns homework in Smartwork5, access your assignments at digital.wwnorton.com/astro5.

digital.wwnorton.com/astro5

In this Exploration, you will see how the H^- thermostat works in the formation of stars. You will need about 20 coins (they do not have to be all the same type).

Place your coins on a sheet of paper and draw a circle around them—the smallest possible circle that will fit all the coins. Then divide the circle into three parts as shown in **Figure 15.26**. This circle represents a star with a changing temperature. The coins represent H^- ions. Removing a coin from the circle means that the H^- ion has turned into a neutral hydrogen atom. Placing a coin in the circle means that the neutral hydrogen atom has become an H^- ion.

Place all the coins back on the circle.

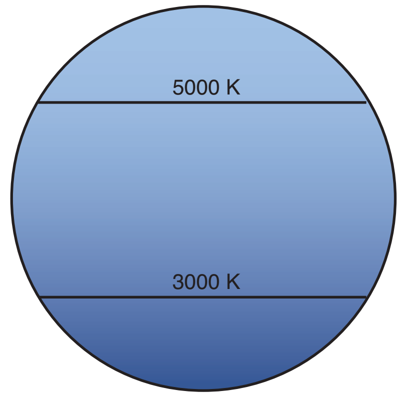


Figure 15.26

1 How many "H^- ions" are now in the star?

The "blanket" of H^- ions holds heat in the star, so the star begins to heat up until it reaches about 5000 K. At that surface temperature, the H^- ions begin to be destroyed. Now that the star is hot, begin removing coins one at a time, starting from the top of the circle and working downward. When you see the line marking 3000 K, stop removing coins.

2 How many " H^- ions" are now in the star?

3 What will happen to the surface temperature of the star, now that there are fewer ions?

When the star cools off to about 3000 K, H^- ions begin to form. Place the coins back on the circle, starting from the bottom and working your way up to the line at 5000 K.

4 How many " H^- ions" are now in the star?

5 What will happen to the surface temperature of the star, now that there are more ions?

Now that the star is hot, begin removing coins one at a time, starting from the top of the circle and working downward. When you see the line marking 3000 K, stop removing coins.

6 What should happen next?

7 Make a circular flowchart that includes the following steps, in the proper order: the star heats up; the star cools down; H^- is formed; H^- is destroyed.

16 Evolution of Low-Mass Stars

Within its core, the Sun fuses more than 4 billion kilograms of hydrogen to helium each second, and although the Sun may seem immortal by human standards, eventually it will run out of fuel. When it does, some 5 billion years from now, the Sun's time on the main sequence will come to an end. In this chapter, we look at how the mass of a star affects the length of its life. Then we examine what happens when a low-mass star like the Sun nears the end of its life.

LEARNING GOALS

By the conclusion of this chapter, you should be able to:

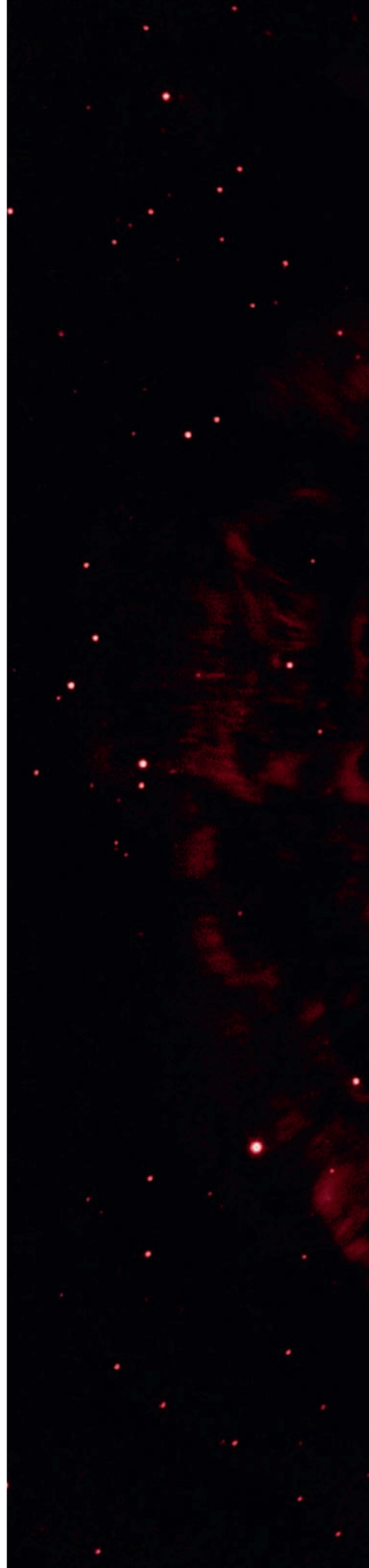
LG 1 Estimate the main-sequence lifetime of a star from its mass.

LG 2 Explain why low-mass stars grow larger and more luminous as they run out of fuel.

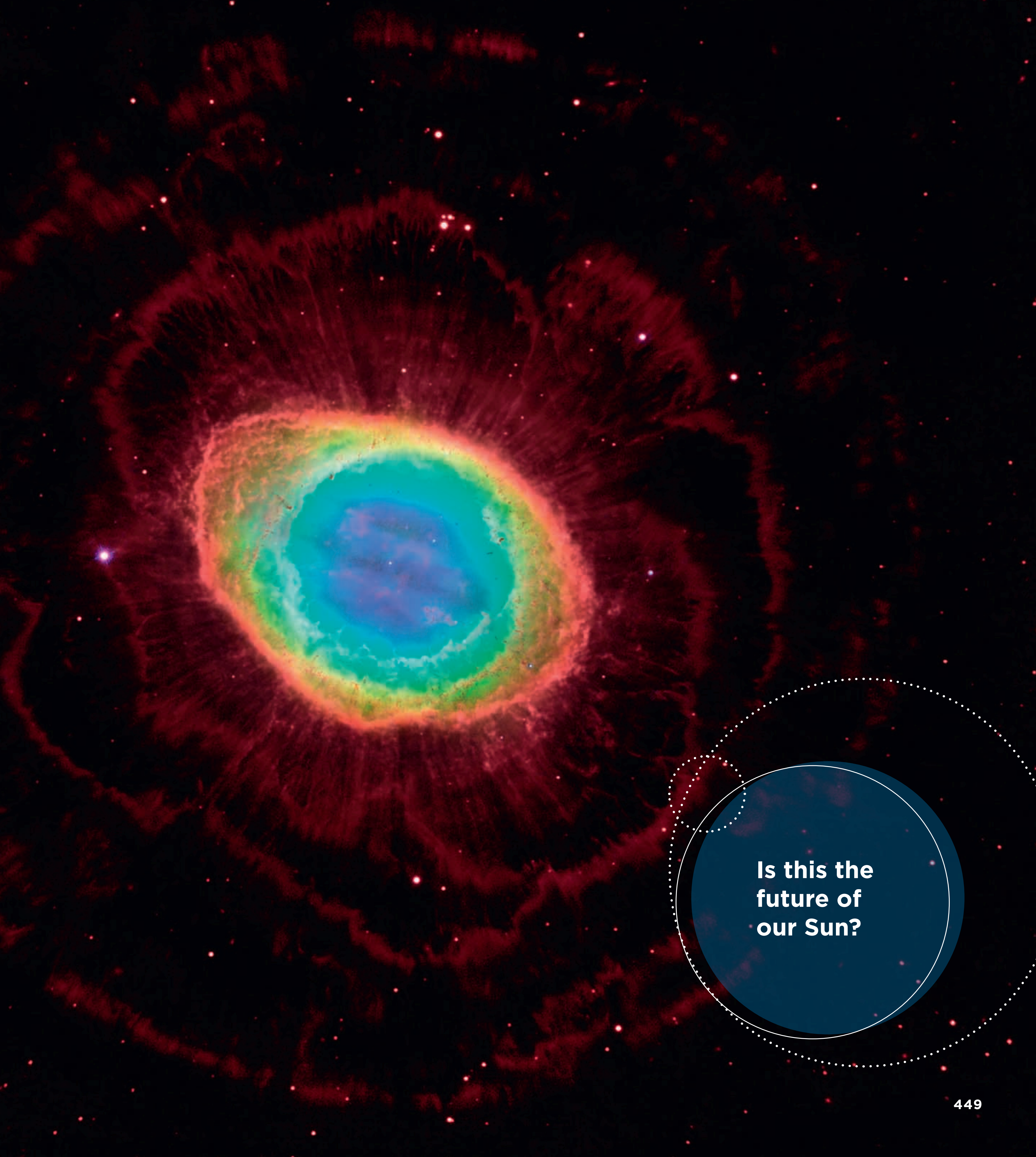
LG 3 Sketch post-main-sequence evolutionary tracks on a Hertzsprung-Russell (H-R) diagram, and list the stages of evolution for low-mass stars.

LG 4 Describe how planetary nebulae and white dwarfs form.

LG 5 Explain how some close binary systems evolve differently than single stars.



When low-mass stars die, they leave behind an expanding nebula of gas and dust as seen in the Hubble Space Telescope image of the Ring Nebula. ▶▶▶



Is this the future of our Sun?

16.1 The Life of a Main-Sequence Star Depends on Its Mass

A star cannot remain on the main sequence forever. It will eventually exhaust the hydrogen fuel in its core, and when it does, its structure will begin to change dramatically. Just as the balance between pressure and gravity within a protostar constantly changes as it evolves toward the main sequence, the balance between pressure and gravity changes as a star evolves beyond the main sequence. The mass and the composition of a star determine the star's life on the main sequence, and these two qualities remain at center stage as the star begins to die. In this section, we explore the changes in the star that trigger its departure from the main sequence.

Observing Stellar Evolution

Suppose you had 1 minute to observe all of the people in a crowded stadium and, from that minute of observation, draw a conclusion about the life cycle of humans. It is possible, but highly unlikely, that you might observe a significant life change such as a birth or a death. More probably, you would observe people of different ages and note some properties of people of various ages that would indicate that some people are young, some are old, and most are in between. But you wouldn't see individual people change over the course of the minute, because a minute is a very small fraction of a typical human lifetime.

Similarly, astronomers observe only a very brief fraction of each star's lifetime. They would have to observe a star like the Sun for several hundred years to be watching it for the equivalent of 1 minute in a human life span. Thus, astronomers do not see individual stars age. But they can observe many, many stars at each stage to piece together an evolutionary picture. Sometimes, just by chance, astronomers observe a star undergoing a dramatic change.

To understand what is happening in the core of the star, theorists use the most powerful computers available to model the nuclear reactions that take place there. These models make predictions about how a star of a given mass and chemical composition will change over its lifetime. These predictions are compared to what astronomers observe. The study of stellar evolution involves back-and-forth between observation and theory, which has led to a general understanding of how and when stars die.

On the basis of this type of analysis, astronomers have found that relatively minor differences in the masses and chemical compositions of two stars can sometimes result in significant, and possibly even dramatic, differences in their fates. Nevertheless, stars can be divided roughly into four broad categories whose members evolve in qualitatively different ways. Stars of spectral types O and B are hot and luminous. These stars have masses above 8 M_{Sun} and are called **high-mass stars**. Other B stars with masses between 3 and 8 M_{Sun} are **medium-mass** stars. Stars of spectral types A, F, G and K, which have masses between 0.5 and 3 M_{Sun}, are considered **low-mass stars** and are typified by the Sun. M stars with masses less than 0.5 M_{Sun} are considered very low-mass stars. In this chapter, we are primarily concerned with the more common stars of mass less than 8 M_{Sun}. In the next chapter, we will consider the fate of more massive stars.

Main-Sequence Lifetime

Think about how long you can drive your car before it runs out of gas: it depends on how much gas your tank holds and on the size and efficiency of your engine. The amount of time your car will run is determined by a competition between these two effects. An SUV uses gas faster than a subcompact, and so it might run out of gas faster, even though it has a much larger gas tank.

The competition between these two effects—tank size and engine size—is most readily expressed as a ratio. How long your car runs is given by the amount of gas in the tank, divided by how quickly the car uses it:

$$\text{Lifetime of tank of gas (hours)} = \frac{\text{Amount of fuel (gallons)}}{\text{Rate at which fuel is used (gallons/hour)}}$$

For example, if you have a 15-gallon tank and your engine is burning fuel at a rate of 3 gallons each hour, then your car will use up all of the gas in 5 hours.

The same principle works for main-sequence stars. The mass of the star determines how much fuel is available. The more massive the star, the more hydrogen available to power nuclear burning. The luminosity of the star indicates the rate at which fuel is used: energy is radiated into space from the surface of a main-sequence star at the same rate at which energy is being generated in its core. If one main-sequence star has twice the luminosity of another, then it must be burning hydrogen at twice the rate of the other.

The **main-sequence lifetime** of a star is the amount of time that it spends on the main sequence, burning hydrogen as its primary source of energy. An expression for the main-sequence lifetime of the star looks very similar to the expression for the time it takes your car to run out of fuel:

$$\text{Lifetime of star} = \frac{\text{Amount of fuel } (\propto \text{ mass of star})}{\text{Rate fuel is used } (\propto \text{ luminosity of star})}$$

The graph in **Figure 16.1** shows that as the mass of a main-sequence star increases, so does the luminosity: this occurs because the mass of the star governs the rate at which nuclear reactions occur in the core. More mass means stronger gravity; stronger gravity means higher pressure in the interior. This higher pressure compresses the core, so more atomic nuclei are packed together into a smaller volume, and it is more likely that they will run into each other and fuse. The increased pressure increases the rate of fuel use. Stronger gravity also increases the temperature in the core and speeds up the atomic nuclei. The nuclei collide more violently, increasing the chances that they will overcome the electric repulsion that pushes the positively charged nuclei apart. The increased temperature increases the rate of fuel use. As a result of the combined effects of temperature and pressure, modest increases in mass can sometimes lead to dramatic increases in the amount of energy released by nuclear burning. Stars with higher masses live shorter lives, not longer ones, because they burn their fuel faster. **Table 16.1** shows the main-sequence lifetimes for stars of different spectral types and masses. This concept is developed further in **Working It Out 16.1**.

Changes in Structure

As discussed in Chapter 14, at the end of the proton-proton chain in main-sequence stars, two ^{3}He nuclei fuse together to form ^{4}He (and two ^{1}H). However,

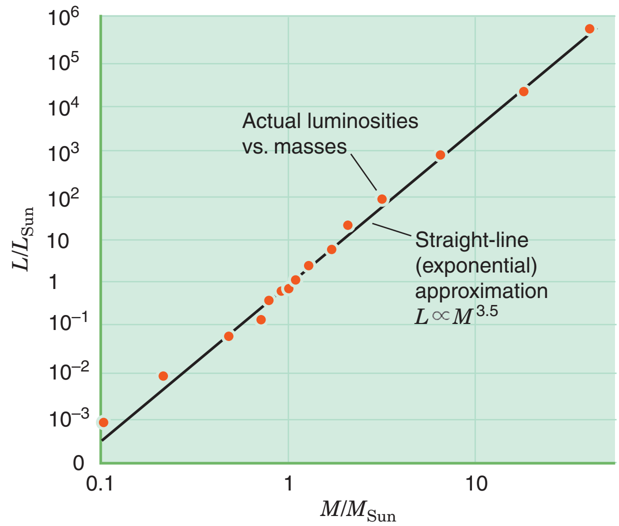


Figure 16.1 This graph plots the mass-luminosity relationship for main-sequence stars: $L \propto M^{3.5}$. The exponent can vary from 2.5 to 5.0, depending on the mass of the star. The average value, over the wide range of main-sequence star masses, is 3.5. Observational data show that the deviation of stars from the average relationship depends on their composition.

TABLE 16.1 Main-Sequence Lifetimes

Spectral Type	Mass (M_{Sun})	Luminosity (L_{Sun})	Main-Sequence Lifetime (years)
O5	60	500,000	3.6×10^5
B0	17.5	32,500	7.8×10^6
B5	5.9	480	1.2×10^8
A0	2.9	39	7×10^8
A5	2.0	12.3	1.8×10^9
F0	1.6	5.2	3.1×10^9
F5	1.4	2.6	4.3×10^9
G0	1.05	1.25	8.9×10^9
G2 (Sun)	1.0	1.0	1.0×10^{10}
G5	0.92	0.8	1.2×10^{10}
K0	0.79	0.55	1.8×10^{10}
K5	0.67	0.32	2.7×10^{10}
M0	0.51	0.08	5.4×10^{10}
M5	0.14	0.008	4.9×10^{11}
M8	~0.08	0.0003	1.1×10^{12}

16.1 Working It Out Estimating Main-Sequence Lifetimes

Astronomers can estimate the lifetime of main-sequence stars either observationally or by modeling the evolution of stars of a given composition, assuming a constant luminosity. Using what is known about how much hydrogen must be converted into helium each second to produce a given amount of energy, as well as the fraction of its hydrogen that a star burns, we can state that the main-sequence lifetime of a star, Lifetime_{MS}, can be expressed as:

$$\text{Lifetime}_{MS} \propto \frac{M_{MS}}{L_{MS}}$$

where M is mass (the amount of fuel), and L is luminosity (the rate the fuel is used). The same equation would apply for the Sun:

$$\text{Lifetime}_{Sun} \propto \frac{M_{Sun}}{L_{Sun}}$$

We can express the lifetime as a ratio, adding in that the computed lifetime of a 1-M_{Sun} star like the Sun is 10 billion (1.0×10^{10}) years:

$$\frac{\text{Lifetime}_{MS}}{\text{Lifetime}_{Sun}} = \frac{\text{Lifetime}_{MS}}{10^{10}\text{ yr}} = \frac{M_{MS}/L_{MS}}{M_{Sun}/L_{Sun}}$$

Multiplying through by 10^{10} years and rearranging the fractions yields

$$\text{Lifetime}_{MS} = 10^{10}\text{ yr} \times \frac{M_{MS}/L_{MS}}{M_{Sun}/L_{Sun}} = 10^{10} \times \frac{M_{MS}/M_{Sun}}{L_{MS}/L_{Sun}}\text{ yr}$$

Now let's compare the lifetime of a star with that of the Sun. The relationship between the mass and the luminosity of stars is such that relatively small differences in the masses of stars result in large differences in their main-sequence luminosities. Figure 16.1 shows the **mass-luminosity relationship**, $L \propto M^{3.5}$, for main-sequence stars. As above, we can express this relationship relative to the Sun's mass and luminosity:

$$\frac{L_{MS}}{L_{Sun}} = \left(\frac{M_{MS}}{M_{Sun}}\right)^{3.5}$$

Substituting the mass-luminosity relationship into the lifetime equation gives us

$$\text{Lifetime}_{MS} = 10^{10} \times \frac{M_{MS}/M_{Sun}}{(M_{MS}/M_{Sun})^{3.5}}\text{ yr} = 10^{10} \times \left(\frac{M_{MS}}{M_{Sun}}\right)^{-2.5}\text{ yr}$$

For example, let's look at a K5 main-sequence star. According to Table 16.1, a K5 star has a mass that is equal to about 0.67 M_{Sun}:

$$\text{Lifetime}_{K5} = 10^{10} \times (0.67)^{-2.5}\text{ yr} = 2.7 \times 10^{10}\text{ yr}$$

Instead of the 10-billion-year life span of the Sun, a K5 star has a main-sequence lifetime 2.7 times larger than the Sun's. Even though the K5 star starts out with less fuel than the Sun, it burns that fuel more slowly, so it lives longer.

at the temperatures found at the centers of main-sequence stars, collisions are not energetic or frequent enough for ^{4}He nuclei to overcome the electric repulsion and fuse into more massive elements.

Hydrogen burns most rapidly at the center of a main-sequence star because the temperature and pressure are highest there. Thus, helium, the product of hydrogen burning, accumulates most rapidly at the center of the star. If we could cut open a star and watch it evolve, we would see its chemical composition changing most rapidly at its center and less rapidly as we move outward. **Figure 16.2** shows how the chemical composition inside a star like the Sun changes throughout its main-sequence lifetime. When the Sun formed, it had a uniform composition of about 70 percent hydrogen and 30 percent helium by mass (Figure 16.2a). Since then, the Sun has produced its energy by converting hydrogen into helium via the proton-proton chain. As hydrogen fused into helium, the helium fraction in the core of the Sun climbed. Today, roughly 5 billion years later, only about 35 percent of the mass in the core of the Sun is hydrogen (Figure 16.2b). Five billion years from now, as the Sun begins to leave the main sequence, the very center of the core will have no hydrogen left to burn (Figure 16.2c).

A main-sequence star slowly changes in response to these changes in structure. As a main-sequence star converts the fuel in its core from hydrogen to helium, its structure must continually shift in response to the changing core composition to maintain energy balance. Between the time the Sun was born and the time it will leave the main sequence, its luminosity will roughly double, with most of this change occurring during the last billion years of its life on the main sequence. Main-sequence evolution is slow and modest in comparison with the events that follow the departure of the star from the main sequence.

CHECK YOUR UNDERSTANDING 16.1

Why does mass determine the main-sequence lifetime of a star? (a) Because more massive stars burn fuel faster and therefore have shorter lives. (b) Because more massive stars have more fuel and therefore have longer lives. (c) Because more massive stars burn different fuels and therefore have longer lives. (d) Because more massive stars have different initial compositions and therefore have shorter lives.

16.2 The Star Leaves the Main Sequence

Once the star has used up the hydrogen in the innermost core, thermal energy leaks out of the helium core into the surrounding layers of the star, but no more energy is generated within the core to replace it. The balance that has maintained the structure of the star throughout its life is now broken. The star's life on the main sequence has come to an end, and its further evolution depends on core temperature changes, which govern fusion reactions.

Electron-Degenerate Matter in the Helium Core

At the enormous internal temperatures of a star, almost all of the electrons have been stripped away from their nuclei by energetic collisions. In other words, the gas is completely ionized—a mixture of electrons and atomic nuclei all flying about freely. Because the size of atomic nuclei is very much less than the distance between nuclei in this gas, most of the space inside the star is empty, with electrons and atomic nuclei filling only a tiny fraction of the star's volume. However, once a low-mass star like the Sun exhausts the hydrogen at its center, the situation changes. Because the star is no longer generating energy to keep it from collapsing, gravity begins to "win," and the helium core begins to collapse and become denser. But because only one electron can occupy a single state at a single time, there is a limit to how dense the core can get. As the matter in the core of the star is compressed further and further, it finally reaches this limit. The electrons are smashed tightly together. This matter at the center of the star is now so dense that a single cubic centimeter has a mass of 1,000 kilograms (kg) or more. Matter in which electrons are packed as closely as possible is called **electron-degenerate** matter.

Electron-degenerate matter has a number of fascinating properties. For example, as more and more helium piles up on the electron-degenerate core, the core *shrinks* in size. This is one of the ways that degenerate matter differs from more normal matter: the more massive it is, the smaller it is. This is noticeably different from normal matter, such as cows: more massive cows are bigger, not smaller. The reason the core shrinks is that the added mass increases the strength of gravity and therefore the weight bearing down on the core, so the electrons are smashed together into a smaller volume. The presence of the electron-degenerate core triggers a chain of events that will dominate the evolution of a 1-M_{Sun} star for the next 50 million years after the hydrogen runs out.

Because the electrons resist being packed more closely together, they produce a type of pressure, known as degeneracy pressure, which pushes outward against gravity. Degeneracy pressure keeps the core from further collapse.

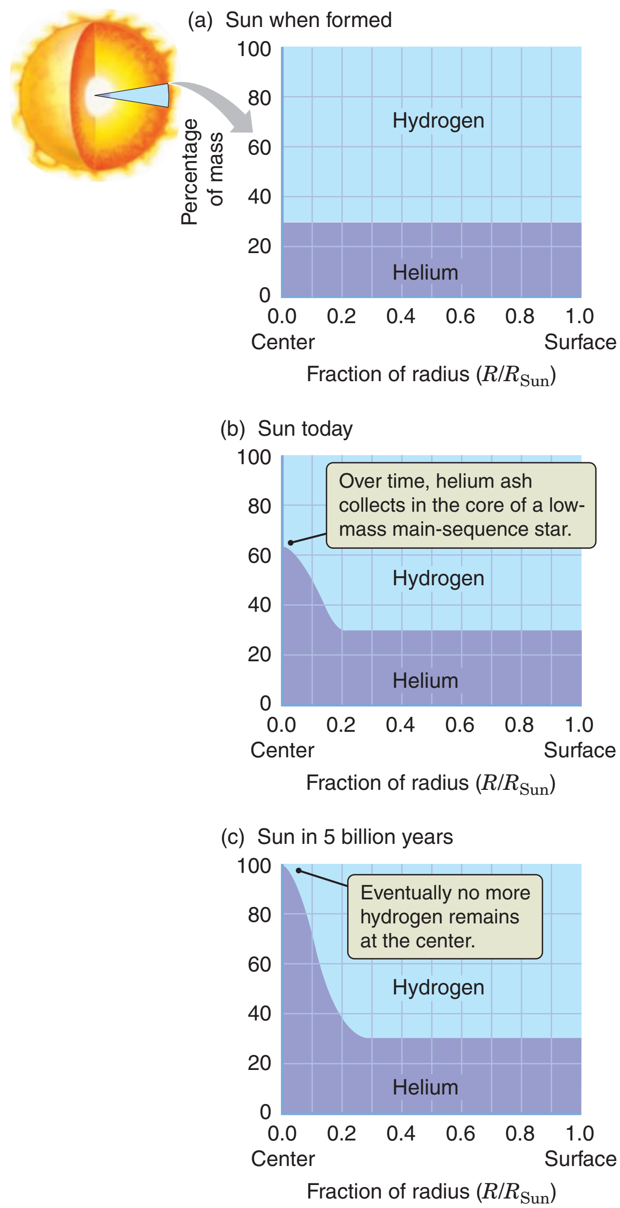


Figure 16.2 Chemical composition of the Sun is plotted here as a percentage of mass against distance from the center of the Sun. (a) When the Sun formed 5 billion years ago, it was evenly mixed: about 30 percent of its mass was helium and 70 percent was hydrogen. (b) Today, the material at the center of the Sun is about 65 percent helium and 35 percent hydrogen. (c) The Sun's main-sequence life will end in about 5 billion years, when all of the hydrogen at its center will be exhausted.

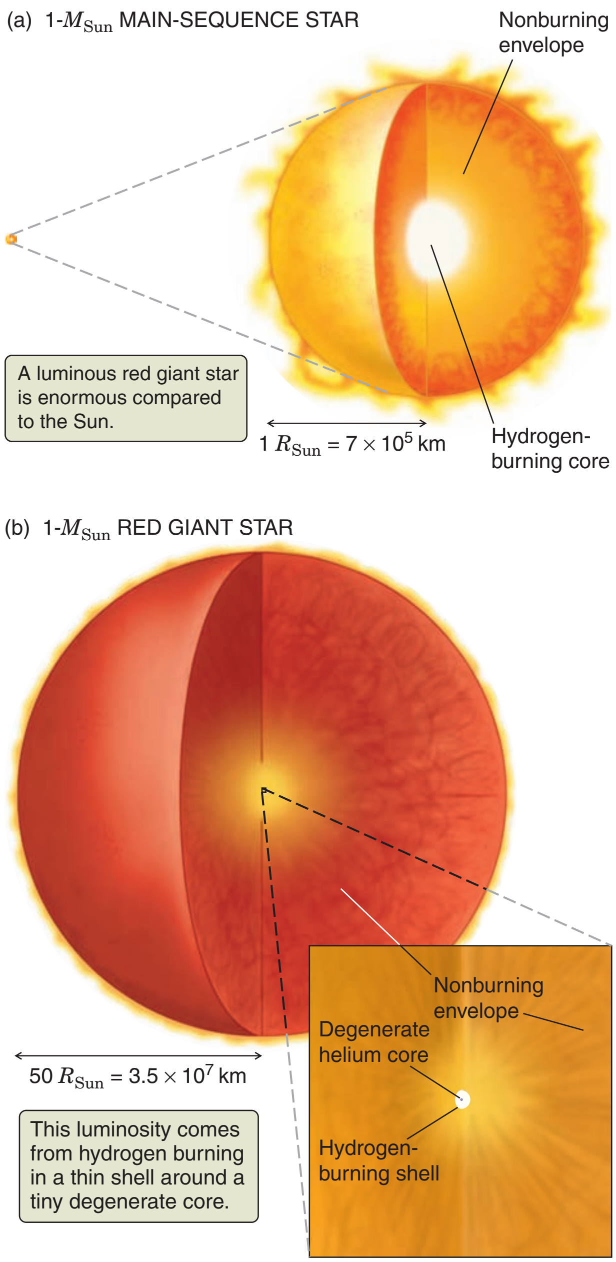


Figure 16.3 The size of the Sun (a) is compared with the size of a star near the top of the red giant branch of the Hertzsprung-Russell (H-R) diagram (b). The structure of the Sun and the core of the red giant are compared in the 50-times larger views identified by the dashed lines.

Hydrogen Shell Burning

After a low-mass star has exhausted the hydrogen at its center, nuclear burning in the core pauses. The layers surrounding the degenerate core still contain hydrogen, and this hydrogen continues to burn. Astronomers call this **hydrogen shell burning** because the hydrogen is burning in a shell surrounding a core of inert helium. This layered structure is like that of a plum, with an internal seed (inert helium), a thin seed coat (hydrogen-burning shell), and a large sphere of flesh (inert hydrogen).

The changes that occur in the heart of a star with a hydrogen-burning shell around a degenerate helium core are reflected in changes in the overall structure of the star. A degenerate core is very compact, so it has very strong gravity, which leads to higher pressure. Faster nuclear reactions in the shell release more energy, so the luminosity of the star increases. With time, the mass of the degenerate helium core grows as more and more hydrogen is converted into helium in the surrounding shell. As the mass of the degenerate helium core grows, so too does its gravitational pull, further increasing the rate of energy generation in the surrounding hydrogen-burning shell. The liberated energy heats the overlying layers of the star, causing them to expand to form a large, luminous giant.

The star becomes larger and more luminous, but also cooler and redder. The enormous surface area of the star allows it to cool very efficiently, so the outer layers become cooler. The relation among radius, temperature, and luminosity that we discussed in Working It Out 13.2 and Chapter 15 still applies ($L = 4\pi\sigma R^2T^4$). In this case, the radius and the luminosity increase, but the temperature surface falls. A red giant star fuses hydrogen in a shell around a degenerate helium core and is larger, more luminous, and redder than it was on the main sequence.

As illustrated in **Figure 16.3**, the internal structure of the main-sequence star (Figure 16.3a) changes as the star evolves to a red giant (Figure 16.3b). The giant can grow to have a luminosity hundreds of times the luminosity of the Sun and a radius of more than 50 solar radii (50 R_{Sun}). Yet at the same time, the core of the giant star is compact: much of the star's mass becomes concentrated into a volume that is only a few times the size of Earth.

The Red Giant Branch

The Hertzsprung-Russell (H-R) diagram shows the changes in a protostar on its way to the main sequence and is also a handy device for keeping track of the star as it evolves away from the main sequence. As soon as the star exhausts the hydrogen in its core, it becomes a **subgiant**: somewhat more luminous, cooler, and larger than it was on the main sequence. Because it grows more luminous but cooler, its position on the H-R diagram travels upward and to the right. As the subgiant continues to evolve, it grows even larger and cooler. When the surface temperature of the subgiant star has dropped by about 1000 kelvins (K) relative to its temperature on the main sequence, H^- ions start to form in great abundance in its atmosphere. Recall from Chapter 15 that H^- ions act as a thermostat, regulating temperature in a protostar by efficiently absorbing and scattering outgoing radiation, trapping this radiation and thus preventing the star from cooling down. The H^- ions serve exactly the same role here as a thermostat regulating how much radiation can escape from the subgiant star and preventing the star from becoming any cooler.

Once the subgiant can cool no further, it becomes a **red giant**—a star that moves almost vertically upward on the H-R diagram as it grows larger and more luminous but remains about the same temperature. A red giant is both redder and larger than the star was on the main sequence. You can think of the path that a star follows on the H-R diagram as it leaves the main sequence as being a tree "branch" growing out of the "trunk" of the main sequence, as shown in **Figure 16.4**. Astronomers call the lower part of this track (which moves somewhat horizontally) the **subgiant branch**. The vertical part is the **red giant branch**. During this time, the luminosity increases, but the mass does not, so the main-sequence mass-luminosity relation no longer applies. The path that a red giant follows on the H-R diagram closely parallels the path that it followed earlier as a collapsing protostar on its way toward the main sequence, except in reverse: this time, the star is moving up that path rather than coming down it. This similarity is not a coincidence. The same physical processes (such as the H^- thermostat) that give rise to the vertical Hayashi track followed by a collapsing protostar also control the relationship of luminosity, size, and surface temperature in an expanding red giant.

As the star leaves the main sequence, the changes in its structure occur slowly at first, but then the star moves up the red giant branch faster and faster. It takes several hundred million years for a star like the Sun to go from the main sequence to the top of the red giant branch. Roughly the first half of this time is spent on the subgiant branch as the star's luminosity increases to about 10 times the luminosity of the Sun (L_{Sun}). During the second half of this time, the helium core of the star grows in mass—but not in radius—as hydrogen is converted to helium in the hydrogen-burning shell and the helium adds to the degenerate core. The increasing mass of the ever more compact helium core increases the force of gravity in the heart of the star. This increased gravity once again increases the pressure and the temperature and thus the rate of nuclear burning, this time in the hydrogen-burning shell. Faster nuclear reactions in the shell convert hydrogen into helium more quickly, so the core grows more rapidly. The star has entered a cycle that feeds on itself. Increasing core mass leads to faster shell burning, and faster shell burning leads to faster core growth. As the core gains mass and the shell becomes more luminous, the outer layers swell. As a result, the star's luminosity climbs at

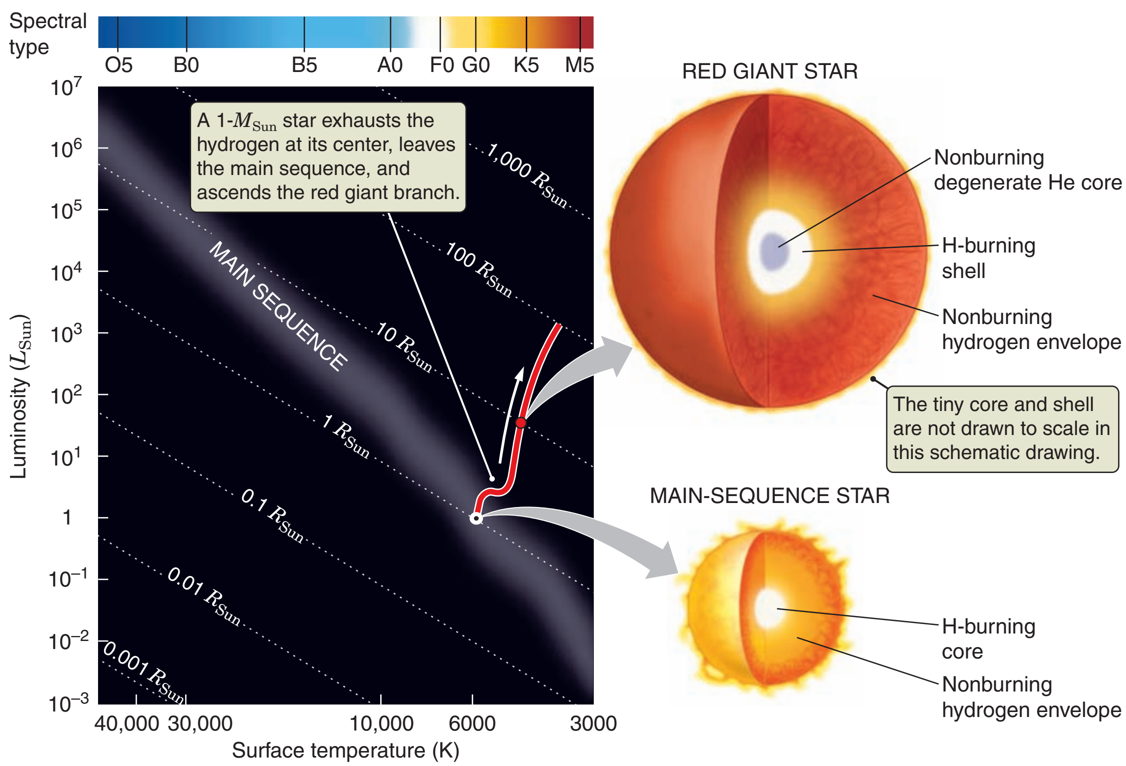


Figure 16.4 A red giant star consists of a degenerate core of helium surrounded by a hydrogen-burning shell. As the star moves up the red giant branch, it comes close to retracing the Hayashi track that it followed when it was a protostar collapsing toward the main sequence.