TEST YOUR UNDERSTANDING OF SECTION 15.5 Four identical strings each carry a sinusoidal wave of frequency 10 Hz. The string tension and wave amplitude are different for different strings. Rank the following strings in order from highest to lowest value of the average wave power: (i) tension 10 N, amplitude 1.0 mm; (ii) tension 40 N, amplitude 1.0 mm; (iii) tension 10 N, amplitude 4.0 mm; (iv) tension 20 N, amplitude 2.0 mm.

ANSWER

(iii), (iv), (ii), (i) Equation (15.25) says that the average power in a sinusoidal wave on a string is $P_{av} = \frac{1}{2}\sqrt{\mu F}\,\omega^2 A^2$. All four strings are identical, so all have the same mass, length, and linear mass density μ. The frequency f is the same for each wave, as is the angular frequency $\omega = 2\pi f$. Hence the average wave power for each string is proportional to the square root of the string tension F and the square of the amplitude A. Compared to string (i), the average power in each string is (ii) $\sqrt{4} = 2$ times greater; (iii) $4^2 = 16$ times greater; and (iv) $\sqrt{2}(2)^2 = 4\sqrt{2}$ times greater.

15.6 WAVE INTERFERENCE, BOUNDARY CONDITIONS, AND SUPERPOSITION

Up to this point we've been discussing waves that propagate continuously in the same direction. But when a wave strikes the boundaries of its medium, all or part of the wave is *reflected*. When you yell at a building wall or a cliff face some distance away, the sound wave is reflected from the rigid surface and you hear an echo. When you flip the end of a rope whose far end is tied to a rigid support, a pulse travels the length of the rope and is reflected back to you. In both cases, the initial and reflected waves overlap in the same region of the medium. We use the term **interference** to refer to what happens when two or more waves pass through the same region at the same time.

As a simple example of wave reflections and the role of the boundary of a wave medium, let's look again at transverse waves on a stretched string. What happens when a wave pulse or a sinusoidal wave arrives at the *end* of the string?

If the end is fastened to a rigid support as in **Fig. 15.18**, it is a *fixed* end that cannot move. The arriving wave exerts a force on the support (drawing 4 in Fig. 15.18); the reaction to this force, exerted by the support on the string, "kicks back" on the string and sets up a reflected pulse or wave traveling in the reverse direction (drawing 7). The reflected pulse moves in the opposite direction from the initial, or *incident,* pulse, and its displacement is also opposite.

The opposite situation from an end that is held stationary is a *free* end, one that is perfectly free to move in the direction perpendicular to the length of the string. For example, the string might be tied to a light ring that slides on a frictionless rod perpendicular to the string, as in **Fig. 15.19**. The ring and rod maintain the tension but exert no transverse force. When a wave arrives at this free end, the ring slides along the rod. The ring reaches a maximum displacement, and both it and the string come momentarily to rest, as in drawing 4 in Fig. 15.19. But the string is now stretched, giving increased tension, so the free end of the string is pulled back down, and again a reflected pulse is produced (drawing 7). As for a fixed end, the reflected pulse moves in the opposite direction from the initial pulse, but now the direction of the displacement is the same as for the initial pulse. The conditions at the end of the string, such as a rigid support or the complete absence of transverse force, are called **boundary conditions.**

The formation of the reflected pulse is similar to the overlap of two pulses traveling in opposite directions. **Figure 15.20** (next page) shows two pulses with the same shape, one inverted with respect to the other, traveling in opposite directions. As the pulses overlap and pass each other, the total displacement of the string is the *algebraic sum* of the displacements at that point in the individual pulses. Because these two pulses have the same shape, the total displacement at point O in the middle of the figure is zero at all times. Thus the motion of the left half of the string would be the same if we cut the string at point O, threw away the right side, and held the end at O fixed. The two pulses on the left side then correspond to the incident and reflected pulses, combining so that the total displacement at O is *always* zero. For this to occur, the reflected pulse must be inverted relative to the incident pulse, just as for reflection from the fixed end in Fig. 15.18.

Figure 15.18 Reflection of a wave pulse at a fixed end of a string. Time increases from top to bottom.

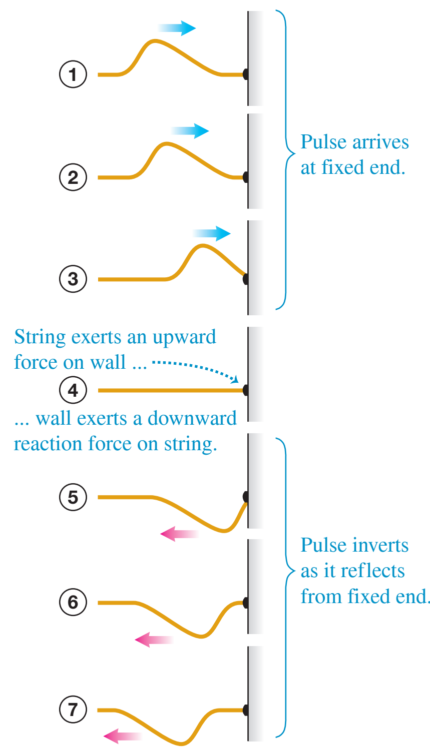


Figure 15.19 Reflection of a wave pulse at a free end of a string. Time increases from top to bottom. (Compare to Fig. 15.18.)

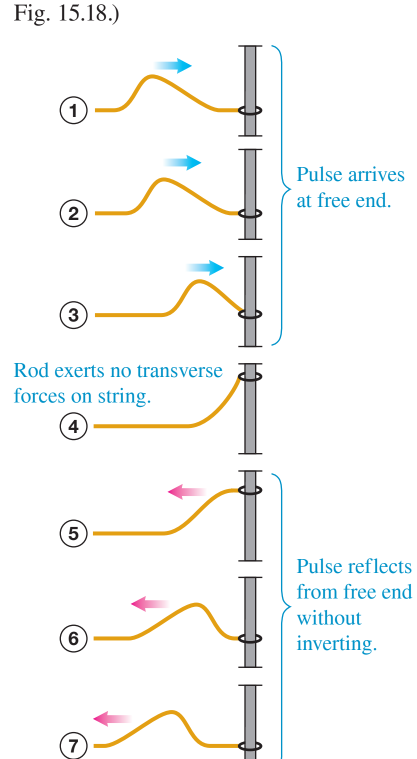


Figure **15.20** Overlap of two wave pulses—one right side up, one inverted—traveling in opposite directions. Time increases from top to bottom.

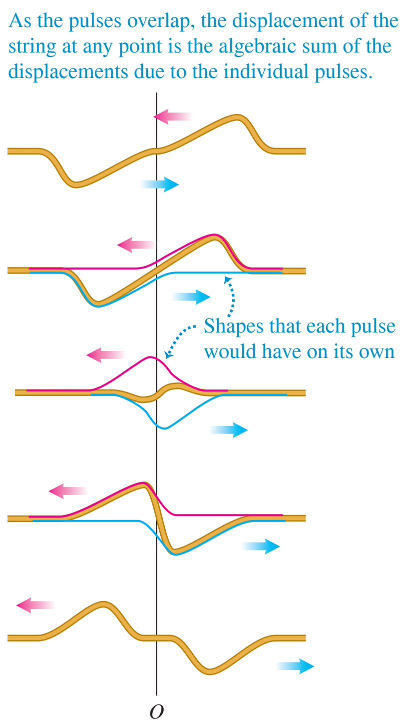


Figure **15.21** Overlap of two wave pulses—both right side up—traveling in opposite directions. Time increases from top to bottom. Compare to Fig. 15.20.

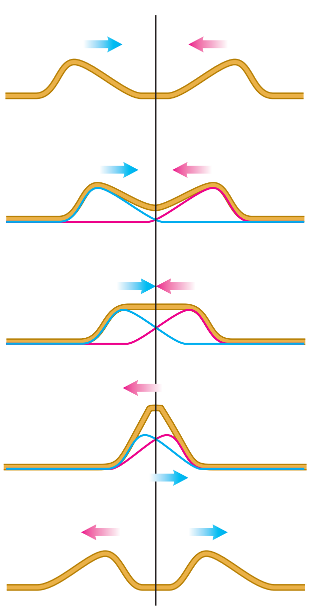


Figure 15.21 shows two pulses with the same shape, traveling in opposite directions but *not* inverted relative to each other. The displacement at point O in the middle of the figure is not zero, but the slope of the string at this point is always zero. According to Eq. (15.20), this corresponds to the absence of any transverse force at this point. In this case the motion of the left half of the string would be the same as if we cut the string at point O and attached the end to a frictionless sliding ring (Fig. 15.19) that maintains tension without exerting any transverse force. In other words, this situation corresponds to reflection of a pulse at a free end of a string at point O. In this case the reflected pulse is *not* inverted.

The Principle of Superposition

Combining the displacements of the separate pulses at each point to obtain the actual displacement is an example of the **principle of superposition:** When two waves overlap, the actual displacement of any point on the string at any time is obtained by adding the displacement the point would have if only the first wave were present and the displacement it would have if only the second wave were present. In other words, the wave function $y(x, t)$ for the resulting motion is obtained by *adding* the two wave functions for the two separate waves:

Principle of superposition:

$$\underbrace{y(x, t)}_{\text{Wave function of combined wave}} = \overbrace{y_1(x, t) + y_2(x, t)}^{\text{Wave functions of two overlapping waves}} \quad (15.27)$$

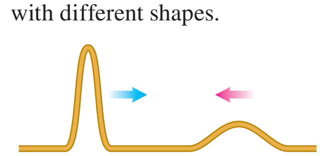
Wave function of combined wave = sum of individual wave functions

Mathematically, this additive property of wave functions follows from the form of the wave equation, Eq. (15.12) or (15.19), which every physically possible wave function must satisfy. Specifically, the wave equation is *linear;* that is, it contains the function $y(x, t)$ only to the first power (there are no terms involving $y(x, t)^2$, $y(x, t)^{1/2}$, etc.). As a result, if any two functions $y_1(x, t)$ and $y_2(x, t)$ satisfy the wave equation separately, their sum $y_1(x, t) + y_2(x, t)$ also satisfies it and is therefore a physically possible motion. Because this principle depends on the linearity of the wave equation and the corresponding linear-combination property of its solutions, it is also called the *principle of linear superposition*. For some physical systems, such as a medium that does not obey Hooke's law, the wave equation is *not* linear; this principle does not hold for such systems.

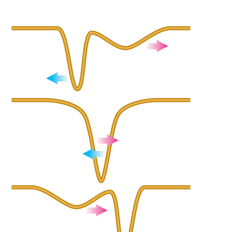
The principle of superposition is of central importance in all types of waves. When a friend talks to you while you are listening to music, you can distinguish the speech and the music from each other. This is precisely because the total sound wave reaching your ears is the algebraic sum of the wave produced by your friend's voice and the wave produced by the speakers of your stereo. If two sound waves did *not* combine in this simple linear way, the sound you would hear in this situation would be a hopeless jumble. Superposition also applies to electromagnetic waves (such as light).

TEST YOUR UNDERSTANDING OF SECTION 15.6
Figure 15.22 shows two wave pulses with different shapes traveling in different directions along a string. Make a series of sketches like Fig. 15.21 showing the shape of the string as the two pulses approach, overlap, and then pass each other.

Figure **15.22** Two wave pulses with different shapes.



ANSWER



15.7 STANDING WAVES ON A STRING

We've looked at the reflection of a wave *pulse* on a string when it arrives at a boundary point (either a fixed end or a free end). Now let's consider what happens when a *sinusoidal* wave on a string is reflected by a fixed end. We'll again approach the problem by considering the superposition of two waves propagating through the string, one representing the incident wave and the other representing the wave reflected at the fixed end.

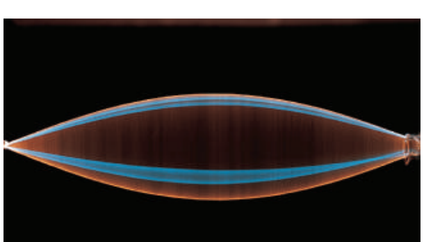
Figure 15.23 shows a string that is fixed at its left end. Its right end is moved up and down in simple harmonic motion to produce a wave that travels to the left; the wave reflected from the fixed end travels to the right. The resulting motion when the two waves combine no longer looks like two waves traveling in opposite directions. The string appears to be subdivided into segments, as in the time-exposure photographs of Figs. 15.23a, 15.23b, 15.23c, and 15.23d. Figure 15.23e shows two instantaneous shapes of the string in Fig. 15.23b. Let's compare this behavior with the waves we studied in Sections 15.1 through 15.5. In a wave that travels along the string, the amplitude is constant and the wave pattern moves with a speed equal to the wave speed. Here, instead, the wave pattern remains in the same position along the string and its amplitude fluctuates. There are particular points called **nodes** (labeled N in Fig. 15.23e) that never move at all. Midway between the nodes are points called **antinodes** (labeled A in Fig. 15.23e) where the amplitude of motion is greatest. Because the wave pattern doesn't appear to be moving in either direction along the string, it is called a **standing wave.** (To emphasize the difference, a wave that *does* move along the string is called a **traveling wave.**)

The principle of superposition explains how the incident and reflected waves combine to form a standing wave. In **Fig. 15.24** (next page) the red curves show a wave traveling to the left. The blue curves show a wave traveling to the right with the same propagation speed, wavelength, and amplitude. The waves are shown at nine instants, $\frac{1}{16}$ of a period apart. At each point along the string, we add the displacements (the values of y) for the two separate waves; the result is the total wave on the string, shown in gold.

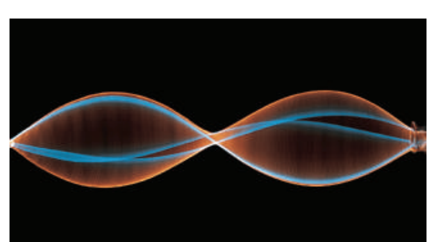
At certain instants, such as $t = \frac{1}{4}T$, the two wave patterns are exactly in phase with each other, and the shape of the string is a sine curve with twice the amplitude of either

Figure **15.23** (a)–(d) Time exposures of standing waves in a stretched string. From (a) to (d), the frequency of oscillation of the right-hand end increases and the wavelength of the standing wave decreases. (e) The extremes of the motion of the standing wave in part (b), with nodes at the center and at the ends. The right-hand end of the string moves very little compared to the antinodes and so is essentially a node.

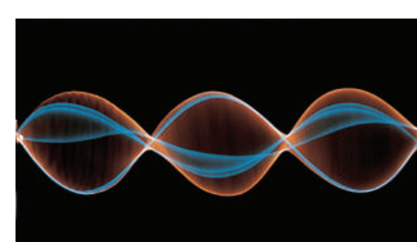
(a) String is one-half wavelength long.



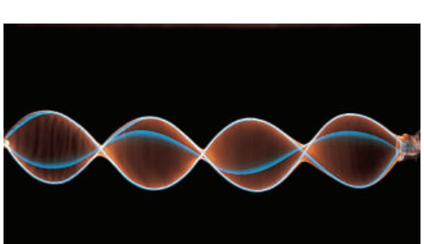
(b) String is one wavelength long.



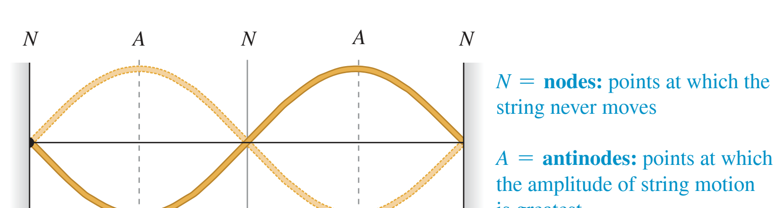
(c) String is one and a half wavelengths long.



(d) String is two wavelengths long.



(e) The shape of the string in (b) at two different instants



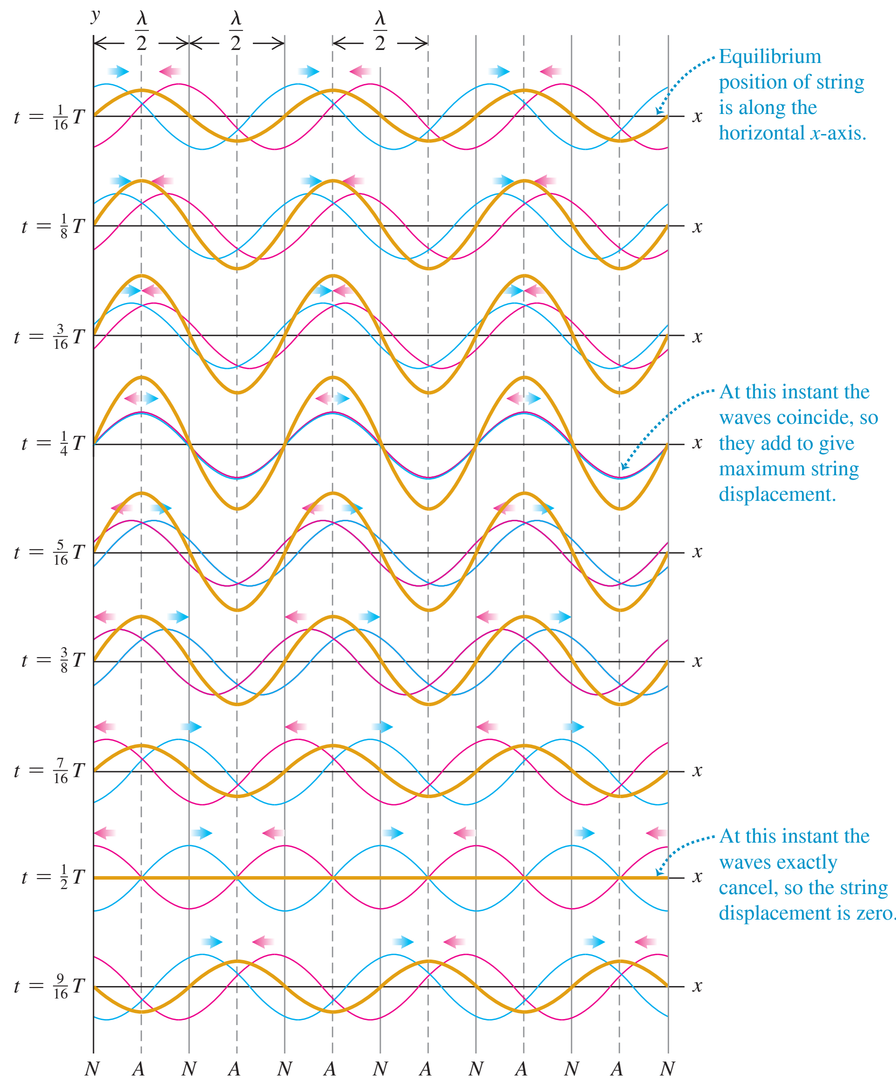


Figure **15.24** Formation of a standing wave. A wave traveling to the left (red curves) combines with a wave traveling to the right (blue curves) to form a standing wave (gold curves).

individual wave. At other instants, such as $t = \frac{1}{2}T$, the two waves are exactly out of phase with each other, and the total wave at that instant is zero. The resultant displacement is *always* zero at those places marked N at the bottom of Fig. 15.24. These are the *nodes.* At a node the displacements of the two waves in red and blue are always equal and opposite and cancel each other out. This cancellation is called **destructive interference.** Midway between the nodes are the points of *greatest* amplitude, or the *antinodes,* marked A. At the antinodes the displacements of the two waves in red and blue are always identical, giving a large resultant displacement; this phenomenon is called **constructive interference.** We can see from the figure that the distance between successive nodes or between successive antinodes is one half-wavelength, or $\lambda/2$.

We can derive a wave function for the standing wave of Fig. 15.24 by adding the wave functions $y_1(x, t)$ and $y_2(x, t)$ for two waves with equal amplitude, period, and wavelength traveling in opposite directions. Here $y_1(x, t)$ (the red curves in Fig. 15.24) represents an incoming, or *incident,* wave traveling to the left along the $+x$-axis, arriving at the point $x = 0$ and being reflected; $y_2(x, t)$ (the blue curves in Fig. 15.24) represents the *reflected* wave traveling to the right from $x = 0$. We noted in Section 15.6 that the wave reflected from a fixed end of a string is inverted, so we give a negative sign to one of the waves:

$$y_1(x, t) = -A\cos(kx + \omega t) \quad \text{(incident wave traveling to the left)}$$

$$y_2(x, t) = A\cos(kx - \omega t) \quad \text{(reflected wave traveling to the right)}$$

The change in sign corresponds to a shift in *phase* of 180° or π radians. At $x = 0$ the motion from the reflected wave is $A\cos\omega t$ and the motion from the incident wave is $-A\cos\omega t$, which we can also write as $A\cos(\omega t + \pi)$. From Eq. (15.27), the wave function for the standing wave is the sum of the individual wave functions:

$$y(x,t) = y_1(x,t) + y_2(x,t) = A[-\cos(kx + \omega t) + \cos(kx - \omega t)]$$

We can rewrite each of the cosine terms by using the identities for the cosine of the sum and difference of two angles: $\cos(a \pm b) = \cos a \cos b \mp \sin a \sin b$. Applying these and combining terms, we obtain the wave function for the standing wave:

$$y(x,t) = y_1(x,t) + y_2(x,t) = (2A\sin kx)\sin\omega t \qquad \text{or}$$

Standing wave on a string, fixed end at $x = 0$:

$$\overset{\text{Wave function}}{y(x,t)} = (\overset{\text{Standing-wave amplitude}}{A_{SW}}\sin \underset{\text{Wave number}}{k}\underset{\text{Position}}{x})\sin \underset{\text{Angular frequency}}{\omega}\overset{\text{Time}}{t} \qquad (15.28)$$

The standing-wave amplitude A_{SW} is twice the amplitude A of either of the original traveling waves: $A_{SW} = 2A$.

Equation (15.28) has two factors: a function of x and a function of t. The factor $A_{SW}\sin kx$ shows that at each instant the shape of the string is a sine curve. But unlike a wave traveling along a string, the wave shape stays in the same position, oscillating up and down as described by the $\sin\omega t$ factor. This behavior is shown by the gold curves in Fig. 15.24. Each point in the string still undergoes simple harmonic motion, but all the points between any successive pair of nodes oscillate *in phase*. This is in contrast to the phase differences between oscillations of adjacent points that we see with a traveling wave.

We can use Eq. (15.28) to find the positions of the nodes; these are the points for which $\sin kx = 0$, so the displacement is *always* zero. This occurs when $kx = 0, \pi, 2\pi, 3\pi, \ldots$, or, using $k = 2\pi/\lambda$,

$$\begin{aligned} x &= 0, \frac{\pi}{k}, \frac{2\pi}{k}, \frac{3\pi}{k}, \ldots \\ &= 0, \frac{\lambda}{2}, \frac{2\lambda}{2}, \frac{3\lambda}{2}, \ldots \end{aligned} \qquad \text{(nodes of a standing wave on a string, fixed end at } x = 0\text{)} \qquad (15.29)$$

In particular, there is a node at $x = 0$, as there should be, since this point is a fixed end of the string.

A standing wave, unlike a traveling wave, *does not* transfer energy from one end to the other. The two waves that form it would individually carry equal amounts of power in opposite directions. There is a local flow of energy from each node to the adjacent antinodes and back, but the *average* rate of energy transfer is zero at every point. If you use the wave function of Eq. (15.28) to evaluate the wave power given by Eq. (15.21), you will find that the average power is zero.

PROBLEM-SOLVING STRATEGY 15.2 Standing Waves

IDENTIFY *the relevant concepts:* Identify the target variables. Then determine whether the problem is purely *kinematic* (involving only such quantities as wave speed v, wavelength λ, and frequency f) or whether *dynamic* properties of the medium (such as F and μ for transverse waves on a string) are also involved.

SET UP *the problem* using the following steps:

1. Sketch the shape of the standing wave at a particular instant. This will help you visualize the nodes (label them N) and antinodes (A). The distance between adjacent nodes (or antinodes) is $\lambda/2$; the distance between a node and the adjacent antinode is $\lambda/4$.
2. Choose the equations you'll use. The wave function for the standing wave, like Eq. (15.28), is often useful.
3. You can determine the wave speed if you know λ and f (or, equivalently, $k = 2\pi/\lambda$ and $\omega = 2\pi f$) or if you know the relevant properties of the medium (for a string, F and μ).

EXECUTE *the solution:* Solve for the target variables. Once you've found the wave function, you can find the displacement y at any point x and at any time t. You can find the velocity and acceleration of a particle in the medium by taking the first and second partial derivatives of y with respect to time.

EVALUATE *your answer:* Compare your numerical answers with your sketch. Check that the wave function satisfies the boundary conditions (for example, the displacement should be zero at a fixed end).

EXAMPLE 15.6 Standing waves on a guitar string — WITH VARIATION PROBLEMS

A guitar string lies along the x-axis when in equilibrium. The end of the string at $x = 0$ (the bridge of the guitar) is fixed. A sinusoidal wave with amplitude $A = 0.750\text{ mm} = 7.50 \times 10^{-4}\text{ m}$ and frequency $f = 440\text{ Hz}$, corresponding to the red curves in Fig. 15.24, travels along the string in the $-x$-direction at 143 m/s. It is reflected from the fixed end, and the superposition of the incident and reflected waves forms a standing wave. (a) Find the equation giving the displacement of a point on the string as a function of position and time. (b) Locate the nodes. (c) Find the amplitude of the standing wave and the maximum transverse velocity and acceleration.

IDENTIFY and SET UP This is a *kinematics* problem (see Problem-Solving Strategy 15.1 in Section 15.3). The target variables are: in part (a), the wave function of the standing wave; in part (b), the locations of the nodes; and in part (c), the maximum displacement y, transverse velocity v_y, and transverse acceleration a_y. Since there is a fixed end at $x = 0$, we can use Eqs. (15.28) and (15.29) to describe this standing wave. We'll need the relationships $\omega = 2\pi f$, $v = \omega/k$, and $v = \lambda f$.

EXECUTE (a) The standing-wave amplitude is $A_{SW} = 2A = 1.50 \times 10^{-3}\text{ m}$ (twice the amplitude of either the incident or reflected wave). The angular frequency and wave number are

$$\omega = 2\pi f = (2\pi\text{ rad})(440\text{ s}^{-1}) = 2760\text{ rad/s}$$

$$k = \frac{\omega}{v} = \frac{2760\text{ rad/s}}{143\text{ m/s}} = 19.3\text{ rad/m}$$

Equation (15.28) then gives

$$\begin{aligned} y(x,t) &= (A_{SW}\sin kx)\sin\omega t \\ &= [(1.50 \times 10^{-3}\text{ m})\sin(19.3\text{ rad/m})x]\sin(2760\text{ rad/s})t \end{aligned}$$

(b) From Eq. (15.29), the positions of the nodes are $x = 0, \lambda/2, \lambda, 3\lambda/2, \ldots$. The wavelength is $\lambda = v/f = (143\text{ m/s})/(440\text{ Hz}) = 0.325\text{ m}$, so the nodes are at $x = 0, 0.163\text{ m}, 0.325\text{ m}, 0.488\text{ m}, \ldots$.

(c) From the expression for $y(x, t)$ in part (a), the maximum displacement from equilibrium is $A_{SW} = 1.50 \times 10^{-3}\text{ m} = 1.50\text{ mm}$. This occurs at the *antinodes,* which are midway between adjacent nodes (that is, at $x = 0.081\text{ m}, 0.244\text{ m}, 0.406\text{ m}, \ldots$).

For a particle on the string at any point x, the transverse (y-) velocity is

$$\begin{aligned} v_y(x,t) &= \frac{\partial y(x,t)}{\partial t} \\ &= [(1.50 \times 10^{-3}\text{ m})\sin(19.3\text{ rad/m})x] \\ &\quad \times [(2760\text{ rad/s})\cos(2760\text{ rad/s})t] \\ &= [(4.15\text{ m/s})\sin(19.3\text{ rad/m})x]\cos(2760\text{ rad/s})t \end{aligned}$$

At an antinode, $\sin(19.3\text{ rad/m})x = \pm 1$ and the transverse velocity varies between $+4.15\text{ m/s}$ and -4.15 m/s. As is always the case in SHM, the maximum velocity occurs when the particle is passing through the equilibrium position ($y = 0$).

The transverse acceleration $a_y(x, t)$ is the *second* partial derivative of $y(x, t)$ with respect to time. You can show that

$$\begin{aligned} a_y(x,t) &= \frac{\partial v_y(x,t)}{\partial t} = \frac{\partial^2 y(x,t)}{\partial t^2} \\ &= [(-1.15 \times 10^4\text{ m/s}^2)\sin(19.3\text{ rad/m})x]\sin(2760\text{ rad/s})t \end{aligned}$$

At the antinodes, the transverse acceleration varies between $+1.15 \times 10^4\text{ m/s}^2$ and $-1.15 \times 10^4\text{ m/s}^2$.

EVALUATE The maximum transverse velocity at an antinode is quite respectable (about 15 km/h, or 9.3 mi/h). But the maximum transverse acceleration is tremendous, 1170 times the acceleration due to gravity! Guitar strings are actually fixed at *both* ends; we'll see the consequences of this in the next section.

KEYCONCEPT A sinusoidal standing wave is the superposition of two sinusoidal waves of the same amplitude and frequency traveling in opposite directions. The wave nodes, which are spaced one half-wavelength apart, are points where the displacement in the standing wave is always zero.

TEST YOUR UNDERSTANDING OF SECTION 15.7 Suppose the frequency of the standing wave in Example 15.6 were doubled from 440 Hz to 880 Hz. Would all of the nodes for $f = 440\text{ Hz}$ also be nodes for $f = 880\text{ Hz}$? If so, would there be additional nodes for $f = 880\text{ Hz}$? If not, which nodes are absent for $f = 880\text{ Hz}$?

ANSWER

Yes, yes Doubling the frequency makes the wavelength half as large. Hence the spacing between nodes (equal to $\lambda/2$) is also half as large. There are nodes at all of the previous positions, but there is also a new node between every pair of old nodes.

15.8 NORMAL MODES OF A STRING

When we described standing waves on a string rigidly held at one end, as in Fig. 15.23, we made no assumptions about the length of the string or about what was happening at the other end. Let's now consider a string of a definite length L, rigidly held at *both* ends. Such strings are found in many musical instruments, including pianos, violins, and guitars. When a guitar string is plucked, a wave is produced in the string; this wave is reflected and re-reflected from the ends of the string, making a standing wave. This standing wave on the string in turn produces a sound wave in the air, with a frequency determined by the properties of the string. This is what makes stringed instruments so useful in making music.

To understand a standing wave on a string fixed at both ends, we first note that the standing wave must have a node at *both* ends of the string. We saw in the preceding section that adjacent nodes are one half-wavelength ($\lambda/2$) apart, so the length of the string must be $\lambda/2$, or $2(\lambda/2)$, or $3(\lambda/2)$, or in general some integer number of half-wavelengths:

$$L = n\frac{\lambda}{2} \quad (n = 1, 2, 3, \ldots) \quad \text{(string fixed at both ends)} \qquad (15.30)$$

That is, if a string with length L is fixed at both ends, a standing wave can exist only if its wavelength satisfies Eq. (15.30).

Solving this equation for λ and labeling the possible values of λ as λ_n, we find

$$\lambda_n = \frac{2L}{n} \quad (n = 1, 2, 3, \ldots) \quad \text{(string fixed at both ends)} \qquad (15.31)$$

Waves can exist on the string if the wavelength is *not* equal to one of these values, but there cannot be a steady wave pattern with nodes and antinodes, and the total wave cannot be a standing wave. Equation (15.31) is illustrated by the standing waves shown in Figs. 15.23a, 15.23b, 15.23c, and 15.23d; these represent $n = 1, 2, 3$, and 4, respectively.

Corresponding to the series of possible standing-wave wavelengths λ_n is a series of possible standing-wave frequencies f_n, each related to its corresponding wavelength by $f_n = v/\lambda_n$. The smallest frequency f_1 corresponds to the largest wavelength (the $n = 1$ case), $\lambda_1 = 2L$:

$$f_1 = \frac{v}{2L} \quad \text{(string fixed at both ends)} \qquad (15.32)$$

This is called the **fundamental frequency.** The other standing-wave frequencies are $f_2 = 2v/2L$, $f_3 = 3v/2L$, and so on. These are all integer multiples of f_1, such as $2f_1, 3f_1, 4f_1$, and so on. We can express *all* the frequencies as

Standing-wave frequencies, string fixed at both ends:

Wave speed — Fundamental frequency $= v/2L$

$$f_n = n\frac{v}{2L} = nf_1 \quad (n = 1, 2, 3, \ldots) \qquad (15.33)$$

Length of string

These frequencies are called **harmonics,** and the series is called a **harmonic series.** Musicians sometimes call f_2, f_3, and so on **overtones;** f_2 is the second harmonic or the first overtone, f_3 is the third harmonic or the second overtone, and so on. The first harmonic is the same as the fundamental frequency (**Fig. 15.25**).

For a string with fixed ends at $x = 0$ and $x = L$, the wave function $y(x, t)$ of the nth standing wave is given by Eq. (15.28) (which satisfies the condition that there is a node at $x = 0$), with $\omega = \omega_n = 2\pi f_n$ and $k = k_n = 2\pi/\lambda_n$:

$$y_n(x, t) = A_{SW} \sin k_n x \sin \omega_n t \qquad (15.34)$$

You can confirm that this wave function has nodes at both $x = 0$ and $x = L$.

A **normal mode** of an oscillating system is a motion in which all particles of the system move sinusoidally with the same frequency. For a system made up of a string of length L fixed at both ends, each of the frequencies given by Eq. (15.33) corresponds to a possible normal-mode pattern. **Figure 15.26** shows the first four normal-mode patterns and their associated frequencies and wavelengths; these correspond to Eq. (15.34) with $n = 1, 2, 3$, and 4. By contrast, a harmonic oscillator, which has only one oscillating particle, has only one normal mode and one characteristic frequency. The string fixed at both ends has infinitely many normal modes ($n = 1, 2, 3, \ldots$) because it is made up of a very large (effectively infinite) number of particles. More complicated oscillating systems also have infinite numbers of normal modes, though with more complex normal-mode patterns (**Fig. 15.27**, next page).

Figure 15.25 Each string of a violin naturally oscillates at its harmonic frequencies, producing sound waves in the air with the same frequencies.

Figure 15.26 The first four normal modes of a string fixed at both ends. (Compare these to the photographs in Fig. 15.23.)

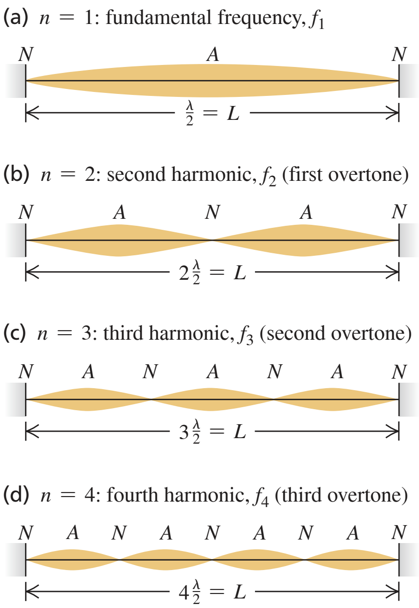
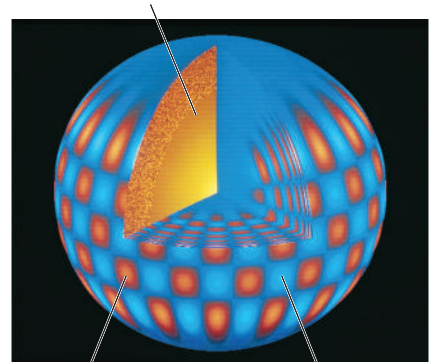


Figure **15.27** Astronomers have discovered that the sun oscillates in several different normal modes. This computer simulation shows one such mode.

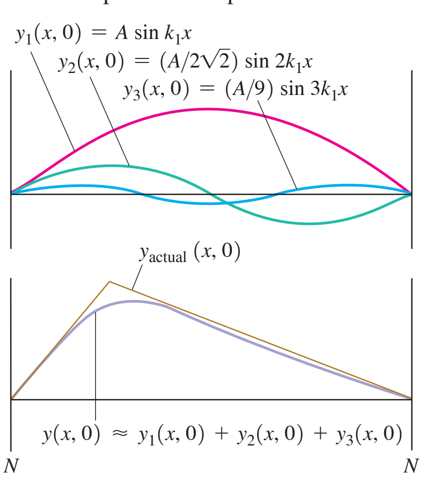


Complex Standing Waves

If we could displace a string so that its shape is the same as one of the normal-mode patterns and then release it, it would vibrate with the frequency of that mode. Such a vibrating string would displace the surrounding air with the same frequency, producing a traveling sinusoidal sound wave that your ears would perceive as a pure tone. But when a string is struck (as in a piano) or plucked (as is done to guitar strings), the shape of the displaced string is *not* one of the patterns in Fig. 15.26. The motion is therefore a combination or *superposition* of many normal modes. Several simple harmonic motions of different frequencies are present simultaneously, and the displacement of any point on the string is the superposition of the displacements associated with the individual modes. The sound produced by the vibrating string is likewise a superposition of traveling sinusoidal sound waves, which you perceive as a rich, complex tone with the fundamental frequency f_1. The standing wave on the string and the traveling sound wave in the air have similar **harmonic content** (the extent to which frequencies higher than the fundamental are present). The harmonic content depends on how the string is initially set into motion. If you pluck the strings of an acoustic guitar in the normal location over the sound hole, the sound that you hear has a different harmonic content than if you pluck the strings next to the fixed end on the guitar body.

It is possible to represent every possible motion of the string as some superposition of normal-mode motions. Finding this representation for a given vibration pattern is called *harmonic analysis*. The sum of sinusoidal functions that represents a complex wave is called a *Fourier series*. **Figure 15.28** shows how a standing wave that is produced by plucking a guitar string of length L at a point $L/4$ from one end can be represented as a combination of sinusoidal functions.

Figure **15.28** When a guitar string is plucked (pulled into a triangular shape) and released, a standing wave results. The standing wave is well represented (except at the sharp maximum point) by the sum of just three sinusoidal functions. Including additional sinusoidal functions further improves the representation.



Standing Waves and String Instruments

From Eq. (15.32), the fundamental frequency of a vibrating string is $f_1 = v/2L$. The speed v of waves on the string is determined by Eq. (15.14), $v = \sqrt{F/\mu}$. Combining these equations, we find

Fundamental frequency, string fixed at both ends

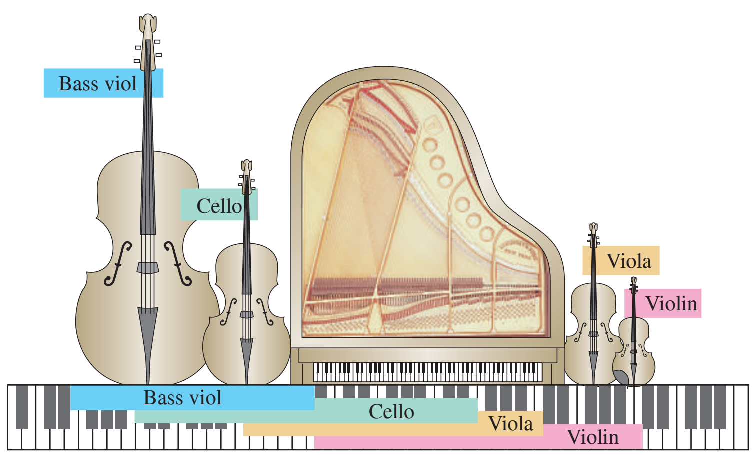
$$f_1 = \frac{1}{2L}\sqrt{\frac{F}{\mu}} \qquad (15.35)$$

Tension in string: F; Mass per unit length: μ; Length of string: L

This is also the fundamental frequency of the sound wave created in the surrounding air by the vibrating string. The inverse dependence of frequency on length L is illustrated by the long strings of the bass (low-frequency) section of the piano or the bass viol compared with the shorter strings of the treble section of the piano or the violin (**Fig. 15.29**). The pitch of a violin or guitar is usually varied by pressing a string against the fingerboard with the fingers to change the length L of the vibrating portion of the string. Increasing the tension F increases the wave speed v and thus increases the frequency (and the pitch). All string instruments are "tuned" to the correct frequencies by varying the tension; you tighten the string to raise the pitch. Finally, increasing the mass per unit length μ decreases the wave speed and thus the frequency. The lower notes on a steel guitar are produced by thicker strings, and one reason for winding the bass strings of a piano with wire is to obtain the desired low frequency from a relatively short string.

Wind instruments such as saxophones and trombones also have normal modes. As for stringed instruments, the frequencies of these normal modes determine the pitch of the musical tones that these instruments produce. We'll discuss these instruments and many other aspects of sound in Chapter 16.

Figure 15.29 Comparing the range of a concert grand piano to the ranges of a bass viol, a cello, a viola, and a violin. In all cases, longer strings produce bass notes and shorter strings produce treble notes.



EXAMPLE 15.7 A giant bass viol

WITH VARIATION PROBLEMS

In an attempt to get your name in *Guinness World Records,* you build a bass viol with strings of length 5.00 m between fixed points. One string, with linear mass density 40.0 g/m, is tuned to a 20.0 Hz fundamental frequency (the lowest frequency that the human ear can hear). Calculate (a) the tension of this string, (b) the frequency and wavelength on the string of the second harmonic, and (c) the frequency and wavelength on the string of the second overtone.

IDENTIFY and SET UP In part (a) the target variable is the string tension F; we'll use Eq. (15.35), which relates F to the known values $f_1 = 20.0\text{ Hz}$, $L = 5.00\text{ m}$, and $\mu = 40.0\text{ g/m}$. In parts (b) and (c) the target variables are the frequency and wavelength of a given harmonic and a given overtone. We determine these from the given length of the string and the fundamental frequency, using Eqs. (15.31) and (15.33).

EXECUTE (a) We solve Eq. (15.35) for F:

$$F = 4\mu L^2 f_1^2 = 4(40.0 \times 10^{-3}\text{ kg/m})(5.00\text{ m})^2(20.0\text{ s}^{-1})^2$$
$$= 1600\text{ N} = 360\text{ lb}$$

(b) From Eqs. (15.33) and (15.31), the frequency and wavelength of the second harmonic $(n = 2)$ are

$$f_2 = 2f_1 = 2(20.0\text{ Hz}) = 40.0\text{ Hz}$$
$$\lambda_2 = \frac{2L}{2} = \frac{2(5.00\text{ m})}{2} = 5.00\text{ m}$$

(c) The second overtone is the "second tone over" (above) the fundamental—that is, $n = 3$. Its frequency and wavelength are

$$f_3 = 3f_1 = 3(20.0\text{ Hz}) = 60.0\text{ Hz}$$
$$\lambda_3 = \frac{2L}{3} = \frac{2(5.00\text{ m})}{3} = 3.33\text{ m}$$

EVALUATE The string tension in a real bass viol is typically a few hundred newtons; the tension in part (a) is a bit higher than that. The wavelengths in parts (b) and (c) are equal to the length of the string and two-thirds the length of the string, respectively, which agrees with the drawings of standing waves in Fig. 15.26.

KEYCONCEPT A string with its ends fixed vibrates in a type of standing wave called a normal mode when all of its particles move sinusoidally with the same frequency. The frequencies of oscillation of the normal modes are integer multiples of a minimum normal-mode frequency, called the fundamental frequency.

EXAMPLE 15.8 From waves on a string to sound waves in air

WITH VARIATION PROBLEMS

What are the frequency and wavelength of the sound waves produced in the air when the string in Example 15.7 is vibrating at its fundamental frequency? The speed of sound in air at 20°C is 344 m/s.

IDENTIFY and SET UP Our target variables are the frequency and wavelength for the *sound wave* produced by the bass viol string. The frequency of the sound wave is the same as the fundamental frequency f_1 of the standing wave, because the string forces the surrounding air to vibrate at the same frequency. The wavelength of the sound wave is $\lambda_{1(\text{sound})} = v_{\text{sound}}/f_1$.

EXECUTE We have $f = f_1 = 20.0$ Hz, so

$$\lambda_{1(\text{sound})} = \frac{v_{\text{sound}}}{f_1} = \frac{344 \text{ m/s}}{20.0 \text{ Hz}} = 17.2 \text{ m}$$

EVALUATE In Example 15.7, the wavelength of the fundamental on the string was $\lambda_{1(\text{string})} = 2L = 2(5.00 \text{ m}) = 10.0 \text{ m}$. Here $\lambda_{1(\text{sound})} = 17.2$ m is greater than that by the factor of $17.2/10.0 = 1.72$. This is as it should be: Because the frequencies of the sound wave and the standing wave are equal, $\lambda = v/f$ says that the wavelengths in air and on the string are in the same ratio as the corresponding wave speeds; here $v_{\text{sound}} = 344$ m/s is greater than $v_{\text{string}} = (10.0 \text{ m})(20.0 \text{ Hz}) = 200$ m/s by just the factor 1.72.

KEYCONCEPT A string vibrating at a certain frequency produces sound waves of the same frequency in the surrounding air. The standing wave on the string and the sound wave will have different wavelengths, however, if the speed of waves on the string does not equal the speed of sound.

TEST YOUR UNDERSTANDING OF SECTION 15.8 While a guitar string is vibrating, you gently touch the midpoint of the string to ensure that the string does not vibrate at that point. Which normal modes *cannot* be present on the string while you are touching it in this way?

ANSWER

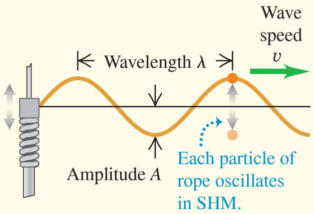
I $n = 1, 3, 5, \ldots$ When you touch the string at its center, you are producing a node at the center. Hence only standing waves with a node at $x = L/2$ are allowed. From Figure 15.26 you can see that the normal modes $n = 1, 3, 5, \ldots$ cannot be present.

CHAPTER 15 SUMMARY

Waves and their properties: A wave is any disturbance that propagates from one region to another. A mechanical wave travels within some material called the medium. The wave speed v depends on the type of wave and the properties of the medium.

In a periodic wave, the motion of each point of the medium is periodic with frequency f and period T. The wavelength λ is the distance over which the wave pattern repeats, and the amplitude A is the maximum displacement of a particle in the medium. The product of λ and f equals the wave speed. A sinusoidal wave is a special periodic wave in which each point moves in simple harmonic motion. (See Example 15.1.)

$$v = \lambda f \tag{15.1}$$



Wave functions and wave dynamics: The wave function $y(x, t)$ describes the displacements of individual particles in the medium. Equations (15.3), (15.4), and (15.7) give the wave equation for a sinusoidal wave traveling in the $+x$-direction. If the wave is moving in the $-x$-direction, the minus signs in the cosine functions are replaced by plus signs. (See Example 15.2.)

The wave function obeys a partial differential equation called the wave equation, Eq. (15.12).

The speed of transverse waves on a string depends on the tension F and mass per unit length μ. (See Example 15.3.)

$$y(x, t) = A\cos\left[\omega\left(\frac{x}{v} - t\right)\right] \tag{15.3}$$

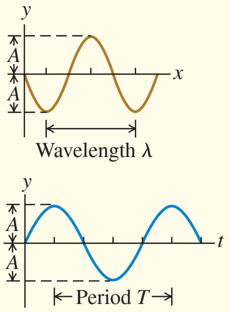
$$y(x, t) = A\cos 2\pi\left[\left(\frac{x}{\lambda} - \frac{t}{T}\right)\right] \tag{15.4}$$

$$y(x, t) = A\cos(kx - \omega t) \tag{15.7}$$

where $k = 2\pi/\lambda$ and $\omega = 2\pi f = vk$

$$\frac{\partial^2 y(x, t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 y(x, t)}{\partial t^2} \tag{15.12}$$

$$v = \sqrt{\frac{F}{\mu}} \text{ (waves on a string)} \tag{15.14}$$



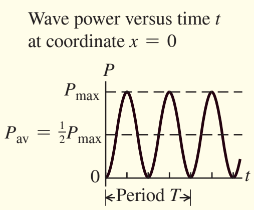
Wave power: Wave motion conveys energy from one region to another. For a sinusoidal mechanical wave, the average power P_{av} is proportional to the square of the wave amplitude and the square of the frequency. For waves that spread out in three dimensions, the wave intensity I is inversely proportional to the square of the distance from the source. (See Examples 15.4 and 15.5.)

$$P_{av} = \tfrac{1}{2}\sqrt{\mu F}\,\omega^2 A^2 \tag{15.25}$$

(average power, sinusoidal wave)

$$\frac{I_1}{I_2} = \frac{r_2^2}{r_1^2} \tag{15.26}$$

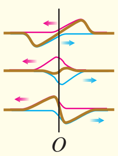
(inverse-square law for intensity)



Wave superposition: A wave reflects when it reaches a boundary of its medium. At any point where two or more waves overlap, the total displacement is the sum of the displacements of the individual waves (principle of superposition).

$$y(x, t) = y_1(x, t) + y_2(x, t) \tag{15.27}$$

(principle of superposition)



Standing waves on a string: When a sinusoidal wave is reflected from a fixed or free end of a stretched string, the incident and reflected waves combine to form a standing sinusoidal wave with nodes and antinodes. Adjacent nodes are spaced a distance $\lambda/2$ apart, as are adjacent antinodes. (See Example 15.6.)

When both ends of a string with length L are held fixed, standing waves can occur only when L is an integer multiple of $\lambda/2$. Each frequency with its associated vibration pattern is called a normal mode. (See Examples 15.7 and 15.8.)

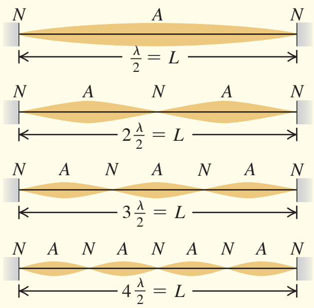
$$y(x, t) = (A_{SW}\sin kx)\sin\omega t \tag{15.28}$$

(standing wave on a string, fixed end at $x = 0$)

$$f_n = n\frac{v}{2L} = nf_1 \quad (n = 1, 2, 3, \dots) \tag{15.33}$$

$$f_1 = \frac{1}{2L}\sqrt{\frac{F}{\mu}} \tag{15.35}$$

(string fixed at both ends)



Chapter 15 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 15.1, 15.2, and 15.3 (Sections 15.2, 15.3, and 15.4) before attempting these problems.

VP15.3.1 A boat is at anchor outside a harbor. A steady sinusoidal ocean wave makes the boat bob up and down with a period of 5.10 s and an amplitude of 1.00 m. The wave has wavelength 30.5 m. For this wave, what are (a) the frequency, (b) the wave speed, (c) the angular frequency, and (d) the wave number?

VP15.3.2 Sound waves in the thin Martian atmosphere travel at 245 m/s. (a) What are the period and wavelength of a 125 Hz sound wave in the Martian atmosphere? (b) What are the frequency and angular frequency of a sound wave in the Martian atmosphere that has wavelength 3.00 m?

VP15.3.3 You are testing a mountain climbing rope that has a linear mass density of 0.0650 kg/m. The rope is held horizontal and is under a tension of 8.00×10^2 N to simulate the stress of supporting a mountain climber's weight. (a) What is the speed of transverse waves on this rope? (b) You oscillate one end of the rope up and down in SHM with frequency 25.0 Hz and amplitude 5.00 mm. What is the wavelength of the resulting waves on the rope? (c) At $t = 0$ the end you are oscillating is at its maximum positive displacement and is instantaneously at rest. Write an equation for the displacement as a function of time at a point 2.50 m from that end. Assume that no wave bounces back from the other end.

VP15.3.4 The tension in a long string is 25.0 N. You oscillate one end of the string up and down with frequency 45.0 Hz. When this end is at its maximum upward displacement, the nearest point that is at its maximum negative displacement is 0.400 m down the string. What are (a) the speed of waves on the string and (b) the linear mass density of the string?

Be sure to review EXAMPLES 15.4 and 15.5 (Section 15.5) before attempting these problems.

VP15.5.1 An athlete exerts a tension of 6.00×10^2 N on one end of a horizontal rope that has length 50.0 m and mass 2.50 kg. The other end is tied to a post. If she wiggles the rope with period 0.575 s and amplitude 3.00 cm, what are (a) the angular frequency of the oscillation and (b) the average rate at which energy is transferred along the rope?

VP15.5.2 A length of piano wire (mass density 5.55×10^{-4} kg/m) is under 185 N of tension. A sinusoidal wave of frequency 256 Hz carries a maximum power of 5.20 W along the wire. What is the amplitude of this wave?

VP15.5.3 A portable audio speaker has a power output of 8.00 W. (a) If the speaker emits sound equally in all directions, what is the sound intensity at a distance of 2.00 m from the speaker? (b) At what distance from the speaker is the intensity equal to 0.045 W/m^2?

VP15.5.4 The "ears" of a frog are two circular membranes located behind the frog's eyes. In one species of frog each membrane is 0.500 cm in radius. If a source of sound has a power output of 2.50×10^{-6} W, emits sound equally in all directions, and is located 1.50 m from the frog, how much sound energy arrives at one of the membranes each second?

Be sure to review EXAMPLES 15.6, 15.7, and 15.8 (Sections 15.7 and 15.8) before attempting these problems.

VP15.8.1 For a standing wave on a string, the distance between nodes is 0.125 m, the frequency is 256 Hz, and the amplitude is 1.40×10^{-3} m. What are (a) the speed of waves on this string, (b) the maximum transverse velocity at an antinode, and (c) the maximum transverse acceleration at an antinode?

VP15.8.2 The G string of a guitar has a fundamental frequency of 196 Hz. The linear mass density of the string is 2.29×10^{-3} kg/m, and the length of string that is free to vibrate (between the nut and bridge of the guitar) is 0.641 m. What are (a) the speed of waves on the G string and (b) the tension in this string?

VP15.8.3 A cable is stretched between two posts 3.00 m apart. The speed of waves on this cable is 96.0 m/s, and the tension in the cable is 175 N. If a standing wave on this cable has five antinodes, what are (a) the wavelength of the standing wave, (b) the frequency of the standing wave, and (c) the linear mass density of the cable?

VP15.8.4 One of the strings on a musical instrument is 0.500 m in length and has linear mass density 1.17×10^{-3} kg/m. The second harmonic on this string has frequency 512 Hz. (a) What is the tension in the string? (b) The speed of sound in air at 20°C is 344 m/s. If the string is vibrating at its fundamental frequency, what is the wavelength of the sound wave that the string produces in air?

BRIDGING PROBLEM Waves on a Rotating Rope

A uniform rope with length L and mass m is held at one end and whirled in a horizontal circle with angular velocity ω. You can ignore the force of gravity on the rope. (a) At a point on the rope a distance r from the end that is held, what is the tension F? (b) What is the speed of transverse waves at this point? (c) Find the time required for a transverse wave to travel from one end of the rope to the other.

SOLUTION GUIDE

IDENTIFY and SET UP

1. Draw a sketch of the situation and label the distances r and L. The tension in the rope will be different at different values of r. Do you see why? Where on the rope do you expect the tension to be greatest? Where do you expect it will be least?
2. Where on the rope do you expect the wave speed to be greatest? Where do you expect it will be least?
3. Think about the portion of the rope that is farther out than r from the end that is held. What forces act on this portion? (Remember that you can ignore gravity.) What is the mass of this portion? How far is its center of mass from the rotation axis?
4. List the unknown quantities and decide which are your target variables.

EXECUTE

5. Draw a free-body diagram for the portion of the rope that is farther out than r from the end that is held.

6. Use your free-body diagram to help you determine the tension in the rope at distance r.

7. Use your result from step 6 to find the wave speed at distance r.

8. Use your result from step 7 to find the time for a wave to travel from one end to the other. (*Hint:* The wave speed is $v = dr/dt$, so the time for the wave to travel a distance dr along the rope is $dt = dr/v$. Integrate this to find the total time. See Appendix B.)

EVALUATE

9. Do your results for parts (a) and (b) agree with your expectations from steps 1 and 2? Are the units correct?

10. Check your result for part (a) by considering the net force on a small segment of the rope at distance r with length dr and mass $dm = (m/L)dr$. [*Hint:* The tension forces on this segment are $F(r)$ on one side and $F(r + dr)$ on the other side. You will get an equation for dF/dr that you can integrate to find F as a function of r.]

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q15.1 Two waves travel on the same string. Is it possible for them to have (a) different frequencies; (b) different wavelengths; (c) different speeds; (d) different amplitudes; (e) the same frequency but different wavelengths? Explain your reasoning.

Q15.2 Under a tension F, it takes 2.00 s for a pulse to travel the length of a taut wire. What tension is required (in terms of F) for the pulse to take 6.00 s instead? Explain how you arrive at your answer.

Q15.3 What kinds of energy are associated with waves on a stretched string? How could you detect such energy experimentally?

Q15.4 The amplitude of a wave decreases gradually as the wave travels along a long, stretched string. What happens to the energy of the wave when this happens?

Q15.5 For the wave motions discussed in this chapter, does the speed of propagation depend on the amplitude? What makes you say this?

Q15.6 The speed of ocean waves depends on the depth of the water; the deeper the water, the faster the wave travels. Use this to explain why ocean waves crest and “break” as they near the shore.

Q15.7 Is it possible to have a longitudinal wave on a stretched string? Why or why not? Is it possible to have a transverse wave on a steel rod? Again, why or why not? If your answer is yes in either case, explain how you would create such a wave.

Q15.8 For transverse waves on a string, is the wave speed the same as the speed of any part of the string? Explain the difference between these two speeds. Which one is constant?

Q15.9 The four strings on a violin have different thicknesses, but are all under approximately the same tension. Do waves travel faster on the thick strings or the thin strings? Why? How does the fundamental vibration frequency compare for the thick versus the thin strings?

Q15.10 A sinusoidal wave can be described by a cosine function, which is negative just as often as positive. So why isn’t the average power delivered by this wave zero?

Q15.11 Two strings of different mass per unit length μ_1 and μ_2 are tied together and stretched with a tension F. A wave travels along the string and passes the discontinuity in μ. Which of the following wave properties will be the same on both sides of the discontinuity, and which will change: speed of the wave; frequency; wavelength? Explain the physical reasoning behind each answer.

Q15.12 A long rope with mass m is suspended from the ceiling and hangs vertically. A wave pulse is produced at the lower end of the rope, and the pulse travels up the rope. Does the speed of the wave pulse change as it moves up the rope, and if so, does it increase or decrease? Explain.

Q15.13 In a transverse wave on a string, the motion of the string is perpendicular to the length of the string. How, then, is it possible for energy to move along the length of the string?

Q15.14 Energy can be transferred along a string by wave motion. However, in a standing wave on a string, no energy can ever be transferred past a node. Why not?

Q15.15 Can a standing wave be produced on a string by superposing two waves traveling in opposite directions with the same frequency but different amplitudes? Why or why not? Can a standing wave be produced by superposing two waves traveling in opposite directions with different frequencies but the same amplitude? Why or why not?

Q15.16 If you stretch a rubber band and pluck it, you hear a (somewhat) musical tone. How does the frequency of this tone change as you stretch the rubber band further? (Try it!) Does this agree with Eq. (15.35) for a string fixed at both ends? Explain.

Q15.17 A musical interval of an *octave* corresponds to a factor of 2 in frequency. By what factor must the tension in a guitar or violin string be increased to raise its pitch one octave? To raise it two octaves? Explain your reasoning. Is there any danger in attempting these changes in pitch?

Q15.18 By touching a string lightly at its center while bowing, a violinist can produce a note exactly one octave above the note to which the string is tuned—that is, a note with exactly twice the frequency. Why is this possible?

Q15.19 As we discussed in Section 15.1, water waves are a combination of longitudinal and transverse waves. Defend the following statement: “When water waves hit a vertical wall, the wall is a node of the longitudinal displacement but an antinode of the transverse displacement.”

Q15.20 Violins are short instruments, while cellos and basses are long. In terms of the frequency of the waves they produce, explain why this is so.

Q15.21 What is the purpose of the frets on a guitar? In terms of the frequency of the vibration of the strings, explain their use.

EXERCISES

Section 15.2 Periodic Waves

15.1 • The speed of sound in air at 20°C is 344 m/s. (a) What is the wavelength of a sound wave with a frequency of 784 Hz, corresponding to the note G_5 on a piano, and how many milliseconds does each vibration take? (b) What is the wavelength of a sound wave one octave higher (twice the frequency) than the note in part (a)?

15.2 • BIO **Ultrasound Imaging.** Sound having frequencies above the range of human hearing (about 20,000 Hz) is called *ultrasound.* Waves above this frequency can be used to penetrate the body and to produce images by reflecting from surfaces. In a typical ultrasound scan, the waves travel through body tissue with a speed of 1500 m/s. For a good, detailed image, the wavelength should be no more than 1.0 mm. What frequency sound is required for a good scan?

15.3 • **Tsunami!** On December 26, 2004, a great earthquake occurred off the coast of Sumatra and triggered immense waves (tsunami) that killed more than 200,000 people. Satellites observing these waves from space measured 800 km from one wave crest to the next and a period between waves of 1.0 hour. What was the speed of these waves in m/s and in km/h? Does your answer help you understand why the waves caused such devastation?

15.4 •• A fisherman notices that his boat is moving up and down periodically, owing to waves on the surface of the water. It takes 2.5 s for the boat to travel from its highest point to its lowest, a total distance of 0.53 m. The fisherman sees that the wave crests are spaced 4.8 m apart. (a) How fast are the waves traveling? (b) What is the amplitude of each wave? (c) If the total vertical distance traveled by the boat were 0.30 m but the other data remained the same, how would the answers to parts (a) and (b) change?

15.5 • BIO (a) **Audible wavelengths.** The range of audible frequencies is from about 20 Hz to 20,000 Hz. What is the range of the wavelengths of audible sound in air? (b) **Visible light.** The range of visible light extends from 380 nm to 750 nm. What is the range of visible frequencies of light? (c) **Brain surgery.** Surgeons can remove brain tumors by using a cavitron ultrasonic surgical aspirator, which produces sound waves of frequency 23 kHz. What is the wavelength of these waves in air? (d) **Sound in the body.** What would be the wavelength of the sound in part (c) in bodily fluids in which the speed of sound is 1480 m/s but the frequency is unchanged?

Section 15.3 Mathematical Description of a Wave

15.6 • A small bead of mass 4.00 g is attached to a horizontal string. Transverse waves of amplitude $A = 0.800$ cm and frequency $f = 20.0$ Hz are set up on the string. Assume the mass of the bead is small enough that the bead doesn't alter the wave motion. During the wave motion, what is the maximum vertical force that the string exerts on the bead?

15.7 • Transverse waves on a string have wave speed 8.00 m/s, amplitude 0.0700 m, and wavelength 0.320 m. The waves travel in the $-x$-direction, and at $t = 0$ the $x = 0$ end of the string has its maximum upward displacement. (a) Find the frequency, period, and wave number of these waves. (b) Write a wave function describing the wave. (c) Find the transverse displacement of a particle at $x = 0.360$ m at time $t = 0.150$ s. (d) How much time must elapse from the instant in part (c) until the particle at $x = 0.360$ m next has maximum upward displacement?

15.8 • A certain transverse wave is described by

$$y(x, t) = (6.50 \text{ mm}) \cos 2\pi\left(\frac{x}{28.0 \text{ cm}} - \frac{t}{0.0360 \text{ s}}\right)$$

Determine the wave's (a) amplitude; (b) wavelength; (c) frequency; (d) speed of propagation; (e) direction of propagation.

15.9 • CALC Which of the following wave functions satisfies the wave equation, Eq. (15.12)? (a) $y(x, t) = A\cos(kx + \omega t)$; (b) $y(x, t) = A\sin(kx + \omega t)$; (c) $y(x, t) = A(\cos kx + \cos \omega t)$. (d) For the wave of part (b), write the equations for the transverse velocity and transverse acceleration of a particle at point x.

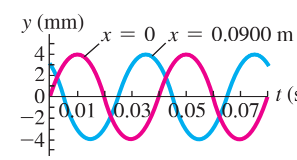
15.10 • A water wave traveling in a straight line on a lake is described by the equation

$$y(x, t) = (2.75 \text{ cm})\cos(0.410 \text{ rad/cm}\, x + 6.20 \text{ rad/s}\, t)$$

where y is the displacement perpendicular to the undisturbed surface of the lake. (a) How much time does it take for one complete wave pattern to go past a fisherman in a boat at anchor, and what horizontal distance does the wave crest travel in that time? (b) What are the wave number and the number of waves per second that pass the fisherman? (c) How fast does a wave crest travel past the fisherman, and what is the maximum speed of his cork floater as the wave causes it to bob up and down?

15.11 • A sinusoidal wave is propagating along a stretched string that lies along the x-axis. The displacement of the string as a function of time is graphed in **Fig. E15.11** for particles at $x = 0$ and at $x = 0.0900$ m. (a) What is the amplitude of the wave? (b) What is the period of the wave? (c) You are told that the two points $x = 0$ and $x = 0.0900$ m are within one wavelength of each other. If the wave is moving in the $+x$-direction, determine the wavelength and the wave speed. (d) If instead the wave is moving in the $-x$-direction, determine the wavelength and the wave speed. (e) Would it be possible to determine definitively the wavelengths in parts (c) and (d) if you were not told that the two points were within one wavelength of each other? Why or why not?

Figure E15.11



15.12 •• CALC **Speed of Propagation vs. Particle Speed.** (a) Show that Eq. (15.3) may be written as

$$y(x, t) = A\cos\left[\frac{2\pi}{\lambda}(x - vt)\right]$$

(b) Use $y(x, t)$ to find an expression for the transverse velocity v_y of a particle in the string on which the wave travels. (c) Find the maximum speed of a particle of the string. Under what circumstances is this equal to the propagation speed v? Less than v? Greater than v?

15.13 •• A transverse wave on a string has amplitude 0.300 cm, wavelength 12.0 cm, and speed 6.00 cm/s. It is represented by $y(x, t)$ as given in Exercise 15.12. (a) At time $t = 0$, compute y at 1.5 cm intervals of x (that is, at $x = 0$, $x = 1.5$ cm, $x = 3.0$ cm, and so on) from $x = 0$ to $x = 12.0$ cm. Graph the results. This is the shape of the string at time $t = 0$. (b) Repeat the calculations for the same values of x at times $t = 0.400$ s and $t = 0.800$ s. Graph the shape of the string at these instants. In what direction is the wave traveling?

Section 15.4 Speed of a Transverse Wave

15.14 • A musical novice learns that doubling the fundamental increases the pitch by one octave and decides to do that to a guitar string. What factor increase in tension would be necessary? Do you think this would be a good idea?

15.15 • One end of a horizontal rope is attached to a prong of an electrically driven tuning fork that vibrates the rope transversely at 120 Hz. The other end passes over a pulley and supports a 1.50 kg mass. The linear mass density of the rope is 0.0480 kg/m. (a) What is the speed of a transverse wave on the rope? (b) What is the wavelength? (c) How would your answers to parts (a) and (b) change if the mass were increased to 3.00 kg?

15.16 • With what tension must a rope with length 2.50 m and mass 0.120 kg be stretched for transverse waves of frequency 40.0 Hz to have a wavelength of 0.750 m?

15.17 •• The upper end of a 3.8-m-long steel wire is fastened to the ceiling, and a 54 kg object is suspended from the lower end of the wire. You observe that it takes a transverse pulse 0.049 s to travel from the bottom to the top of the wire. What is the mass of the wire?

15.18 •• A 1.50 m string of weight 0.0125 N is tied to the ceiling at its upper end, and the lower end supports a weight W. Ignore the very small variation in tension along the length of the string that is produced by the weight of the string. When you pluck the string slightly, the waves traveling up the string obey the equation

$$y(x, t) = (8.50\ \text{mm})\cos(172\ \text{rad/m}\ x - 4830\ \text{rad/s}\ t)$$

Assume that the tension of the string is constant and equal to W. (a) How much time does it take a pulse to travel the full length of the string? (b) What is the weight W? (c) How many wavelengths are on the string at any instant of time? (d) What is the equation for waves traveling *down* the string?

15.19 • A thin, 75.0 cm wire has a mass of 16.5 g. One end is tied to a nail, and the other end is attached to a screw that can be adjusted to vary the tension in the wire. (a) To what tension (in newtons) must you adjust the screw so that a transverse wave of wavelength 3.33 cm makes 625 vibrations per second? (b) How fast would this wave travel?

15.20 •• A heavy rope 6.00 m long and weighing 29.4 N is attached at one end to a ceiling and hangs vertically. A 0.500 kg mass is suspended from the lower end of the rope. What is the speed of transverse waves on the rope at the (a) bottom of the rope, (b) middle of the rope, and (c) top of the rope? (d) Is the tension in the middle of the rope the average of the tensions at the top and bottom of the rope? Is the wave speed at the middle of the rope the average of the wave speeds at the top and bottom? Explain.

Section 15.5 Energy in Wave Motion

15.21 •• In Example 15.4 the average power that Throckmorton puts into the clothesline is small, about 1 W. How much better can you do? Assume a rope that has linear mass density 0.500 kg/m, so 1 m of the rope weighs about 1 lb. The rope is long and is attached to a post at one end. You hold the other end of the rope in your hand and supply sinusoidal wave pulses by moving your arm up and down. Estimate the amplitude of the pulses to be the length of your arm. Estimate the maximum tension you can supply to the rope by pulling on it horizontally, and estimate the time for you to complete each pulse. (a) Ignoring any effects from reflection of the pulses from the other end of the rope, what average power can you supply to the rope? (b) You try to increase your power output by halving the amplitude so you can double the frequency of the pulses. What change in P_{av} does this produce?

15.22 •• A piano wire with mass 3.00 g and length 80.0 cm is stretched with a tension of 25.0 N. A wave with frequency 120.0 Hz and amplitude 1.6 mm travels along the wire. (a) Calculate the average power carried by the wave. (b) What happens to the average power if the wave amplitude is halved?

15.23 • A horizontal wire is stretched with a tension of 94.0 N, and the speed of transverse waves for the wire is 406 m/s. What must the amplitude of a traveling wave of frequency 69.0 Hz be for the average power carried by the wave to be 0.365 W?

15.24 •• **Threshold of Pain.** You are investigating the report of a UFO landing in an isolated portion of New Mexico, and you encounter a strange object that is radiating sound waves uniformly in all directions. Assume that the sound comes from a point source and that you can ignore reflections. You are slowly walking toward the source. When you are 7.5 m from it, you measure its intensity to be 0.11 W/m^2. An intensity of 1.0 W/m^2 is often used as the "threshold of pain." How much closer to the source can you move before the sound intensity reaches this threshold?

15.25 •• A jet plane at takeoff can produce sound of intensity 10.0 W/m^2 at 30.0 m away. But you prefer the tranquil sound of normal conversation, which is 1.0 μW/m^2. Assume that the plane behaves like a point source of sound. (a) What is the closest distance you should live from the airport runway to preserve your peace of mind? (b) What intensity from the jet does your friend experience if she lives twice as far from the runway as you do? (c) What power of sound does the jet produce at takeoff?

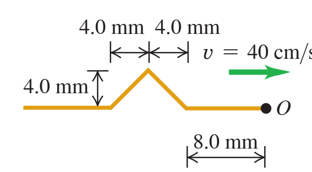
15.26 • A fellow student with a mathematical bent tells you that the wave function of a traveling wave on a thin rope is $y(x, t) = (2.30\ \text{mm})\cos[(6.98\ \text{rad/m})x + (742\ \text{rad/s})t]$. Being more practical, you measure the rope to have a length of 1.35 m and a mass of 0.00338 kg. You are then asked to determine the following: (a) amplitude; (b) frequency; (c) wavelength; (d) wave speed; (e) direction the wave is traveling; (f) tension in the rope; (g) average power transmitted by the wave.

15.27 • **Energy Output.** By measurement you determine that sound waves are spreading out equally in all directions from a point source and that the intensity is 0.026 W/m^2 at a distance of 4.3 m from the source. (a) What is the intensity at a distance of 3.1 m from the source? (b) How much sound energy does the source emit in one hour if its power output remains constant?

Section 15.6 Wave Interference, Boundary Conditions, and Superposition

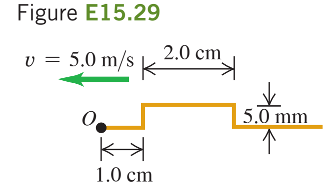
15.28 • **Reflection.** A wave pulse on a string has the dimensions shown in **Fig. E15.28** at $t = 0$. The wave speed is 40 cm/s. (a) If point O is a fixed end, draw the total wave on the string at t = 15 ms, 20 ms, 25 ms, 30 ms, 35 ms, 40 ms, and 45 ms. (b) Repeat part (a) for the case in which point O is a free end.

Figure **E15.28**



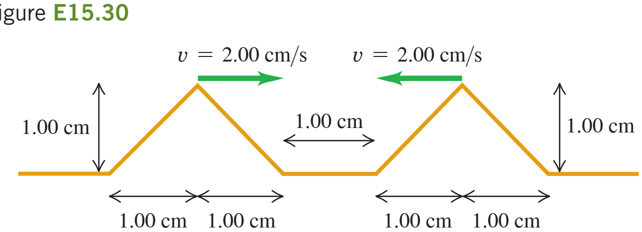
15.29 • **Reflection.** A wave pulse on a string has the dimensions shown in **Fig. E15.29** at $t = 0$. The wave speed is 5.0 m/s. (a) If point O is a fixed end, draw the total wave on the string at t = 1.0 ms, 2.0 ms, 3.0 ms, 4.0 ms, 5.0 ms, 6.0 ms, and 7.0 ms. (b) Repeat part (a) for the case in which point O is a free end.

Figure **E15.29**



15.30 • **Interference of Triangular Pulses.** Two triangular wave pulses are traveling toward each other on a stretched string as shown in **Fig. E15.30**. Each pulse is identical to the other and travels at 2.00 cm/s. The leading edges of the pulses are 1.00 cm apart at $t = 0$. Sketch the shape of the string at t = 0.250 s, t = 0.500 s, t = 0.750 s, t = 1.000 s, and t = 1.250 s.

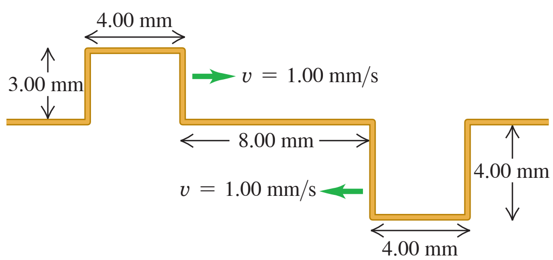
Figure **E15.30**



15.31 • Suppose that the left-traveling pulse in Exercise 15.30 is *below* the level of the unstretched string instead of above it. Make the same sketches that you did in that exercise.

15.32 •• **Interference of Rectangular Pulses. Figure E15.32** shows two rectangular wave pulses on a stretched string traveling toward each other. Each pulse is traveling with a speed of 1.00 mm/s and has the height and width shown in the figure. If the leading edges of the pulses are 8.00 mm apart at $t = 0$, sketch the shape of the string at $t = 4.00$ s, $t = 6.00$ s, and $t = 10.0$ s.

Figure **E15.32**



Section 15.7 Standing Waves on a String

Section 15.8 Normal Modes of a String

15.33 •• For a violin, estimate the length of the portions of the strings that are free to vibrate. (a) The frequency of the note played by the open E5 string vibrating in its fundamental standing wave is 659 Hz. Use your estimate of the length to calculate the wave speed for the transverse waves on the string. (b) The vibrating string produces sound waves in air with the same frequency as that of the string. Use 344 m/s for the speed of sound in air and calculate the wavelength of the E5 note in air. Which is larger: the wavelength on the string or the wavelength in air? (c) Repeat parts (a) and (b) for a bass viol, which is typically played by a person standing up. Start your calculation by estimating the length of the bass viol string that is free to vibrate. The G2 string produces a note with frequency 98 Hz when vibrating in its fundamental standing wave.

15.34 •• **CALC** Adjacent antinodes of a standing wave on a string are 15.0 cm apart. A particle at an antinode oscillates in simple harmonic motion with amplitude 0.850 cm and period 0.0750 s. The string lies along the $+x$-axis and is fixed at $x = 0$. (a) How far apart are the adjacent nodes? (b) What are the wavelength, amplitude, and speed of the two traveling waves that form this pattern? (c) Find the maximum and minimum transverse speeds of a point at an antinode. (d) What is the shortest distance along the string between a node and an antinode?

15.35 • Standing waves on a wire are described by Eq. (15.28), with $A_{SW} = 2.50$ mm, $\omega = 942$ rad/s, and $k = 0.750\pi$ rad/m. The left end of the wire is at $x = 0$. At what distances from the left end are (a) the nodes of the standing wave and (b) the antinodes of the standing wave?

15.36 • A 1.50-m-long rope is stretched between two supports with a tension that makes the speed of transverse waves 62.0 m/s. What are the wavelength and frequency of (a) the fundamental; (b) the second overtone; (c) the fourth harmonic?

15.37 • A wire with mass 40.0 g is stretched so that its ends are tied down at points 80.0 cm apart. The wire vibrates in its fundamental mode with frequency 60.0 Hz and with an amplitude at the antinodes of 0.300 cm. (a) What is the speed of propagation of transverse waves in the wire? (b) Compute the tension in the wire. (c) Find the maximum transverse velocity and acceleration of particles in the wire.

15.38 • A piano tuner stretches a steel piano wire with a tension of 800 N. The steel wire is 0.400 m long and has a mass of 3.00 g. (a) What is the frequency of its fundamental mode of vibration? (b) What is the number of the highest harmonic that could be heard by a person who is capable of hearing frequencies up to 10,000 Hz?

15.39 • **CALC** A thin, taut string tied at both ends and oscillating in its third harmonic has its shape described by the equation $y(x, t) = (5.60\text{ cm})\sin[(0.0340\text{ rad/cm})x]\sin[(50.0\text{ rad/s})t]$, where the origin is at the left end of the string, the x-axis is along the string, and the y-axis is perpendicular to the string. (a) Draw a sketch that shows the standing-wave pattern. (b) Find the amplitude of the two traveling waves that make up this standing wave. (c) What is the length of the string? (d) Find the wavelength, frequency, period, and speed of the traveling waves. (e) Find the maximum transverse speed of a point on the string. (f) What would be the equation $y(x, t)$ for this string if it were vibrating in its eighth harmonic?

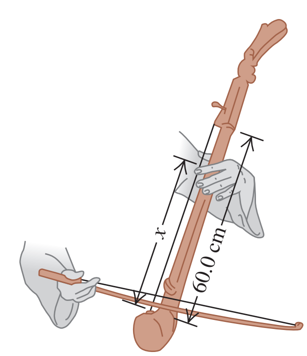
15.40 • The wave function of a standing wave is $y(x, t) = (4.44\text{ mm})\sin[(32.5\text{ rad/m})x]\sin[(754\text{ rad/s})t]$. For the two traveling waves that make up this standing wave, find the (a) amplitude; (b) wavelength; (c) frequency; (d) wave speed; (e) wave functions. (f) From the information given, can you determine which harmonic this is? Explain.

15.41 • Standing waves are produced on a string that is held fixed at both ends. The tension in the string is kept constant. (a) For the second overtone standing wave the node-to-node distance is 8.00 cm. What is the length of the string? (b) What is the node-to-node distance for the fourth harmonic standing wave?

15.42 •• One string of a certain musical instrument is 75.0 cm long and has a mass of 8.75 g. It is being played in a room where the speed of sound is 344 m/s. (a) To what tension must you adjust the string so that, when vibrating in its second overtone, it produces sound of wavelength 0.765 m? (Assume that the breaking stress of the wire is very large and isn't exceeded.) (b) What frequency sound does this string produce in its fundamental mode of vibration?

15.43 • The portion of the string of a certain musical instrument between the bridge and upper end of the finger board (that part of the string that is free to vibrate) is 60.0 cm long, and this length of the string has mass 2.00 g. The string sounds an A_4 note (440 Hz) when played. (a) Where must the player put a finger (what distance x from the bridge) to play a D_5 note (587 Hz)? (See **Fig. E15.43**.) For both the A_4 and D_5 notes, the string vibrates in its fundamental mode. (b) Without retuning, is it possible to play a G_4 note (392 Hz) on this string? Why or why not?

Figure **E15.43**



15.44 • (a) A horizontal string tied at both ends is vibrating in its fundamental mode. The traveling waves have speed v, frequency f, amplitude A, and wavelength λ. Calculate the maximum transverse velocity and maximum transverse acceleration of points located at (i) $x = \lambda/2$, (ii) $x = \lambda/4$, and (iii) $x = \lambda/8$, from the left-hand end of the string. (b) At each of the points in part (a), what is the amplitude of the motion? (c) At each of the points in part (a), how much time does it take the string to go from its largest upward displacement to its largest downward displacement?

PROBLEMS

15.45 •• A sinusoidal wave with wavelength 0.400 m travels along a string. The maximum transverse speed of a point on the string is 3.00 m/s and the maximum transverse acceleration is $8.50 \times 10^4\text{ m/s}^2$. What are the propagation speed v and the amplitude A of the wave?

15.46 • A transverse wave on a rope is given by

$$y(x, t) = (0.750\text{ cm})\cos\pi[(0.400\text{ cm}^{-1})x + (250\text{ s}^{-1})t]$$

(a) Find the amplitude, period, frequency, wavelength, and speed of propagation. (b) Sketch the shape of the rope at these values of t: 0, 0.0005 s, 0.0010 s. (c) Is the wave traveling in the $+x$- or $-x$-direction? (d) The mass per unit length of the rope is 0.0500 kg/m. Find the tension. (e) Find the average power of this wave.

15.47 • CALC A transverse sine wave with an amplitude of 2.50 mm and a wavelength of 1.80 m travels from left to right along a long, horizontal, stretched string with a speed of 36.0 m/s. Take the origin at the left end of the undisturbed string. At time $t = 0$ the left end of the string has its maximum upward displacement. (a) What are the frequency, angular frequency, and wave number of the wave? (b) What is the function $y(x, t)$ that describes the wave? (c) What is $y(t)$ for a particle at the left end of the string? (d) What is $y(t)$ for a particle 1.35 m to the right of the origin? (e) What is the maximum magnitude of transverse velocity of any particle of the string? (f) Find the transverse displacement and the transverse velocity of a particle 1.35 m to the right of the origin at time $t = 0.0625$ s.

15.48 •• CP A 1750 N irregular beam is hanging horizontally by its ends from the ceiling by two vertical wires (A and B), each 1.25 m long and weighing 0.290 N. The center of gravity of this beam is one-third of the way along the beam from the end where wire A is attached. If you pluck both strings at the same time at the beam, what is the time delay between the arrival of the two pulses at the ceiling? Which pulse arrives first? (Ignore the effect of the weight of the wires on the tension in the wires.)

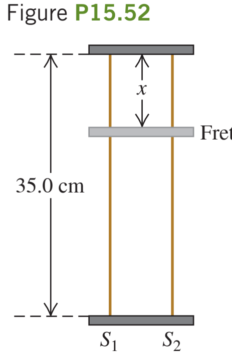
15.49 ••• CP One end of a light uniform rod is attached to a wall by a frictionless hinge. The rod is held in a horizontal position by a wire that runs from the other end of the rod to the wall. The wire has length 2.00 m and makes an angle of 30.0° with the rod. A block with mass m is suspended by a light rope attached to the middle of the rod. The transverse fundamental standing wave on the wire has frequency f. The mass m is varied and for each value the frequency is measured. You plot f^2 versus m and find that your data lie close to a straight line with slope 20.4 $\text{kg}^{-1} \cdot \text{s}^{-2}$. What is the mass of the wire? Assume that the change in the length of the wire when the tension changes is small enough to neglect.

15.50 •• Weightless Ant. An ant with mass m is standing peacefully on top of a horizontal, stretched rope. The rope has mass per unit length μ and is under tension F. Without warning, Cousin Throckmorton starts a sinusoidal transverse wave of wavelength λ propagating along the rope. The motion of the rope is in a vertical plane. What minimum wave amplitude will make the ant become momentarily weightless? Assume that m is so small that the presence of the ant has no effect on the propagation of the wave.

15.51 •• You must determine the length of a long, thin wire that is suspended from the ceiling in the atrium of a tall building. A 2.00-cm-long piece of the wire is left over from its installation. Using an analytical balance, you determine that the mass of the spare piece is 14.5 μg. You then hang a 0.400 kg mass from the lower end of the long, suspended wire. When a small-amplitude transverse wave pulse is sent up that wire, sensors at both ends measure that it takes the wave pulse 26.7 ms to travel the length of the wire. (a) Use these measurements to calculate the length of the wire. Assume that the weight of the wire has a negligible effect on the speed of the transverse waves. (b) Discuss the accuracy of the approximation made in part (a).

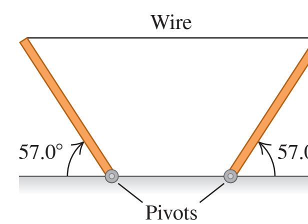
15.52 •• Music. You are designing a two-string instrument with metal strings 35.0 cm long, as shown in **Fig. P15.52**. Both strings are under the *same tension.* String S_1 has a mass of 8.00 g and produces the note middle C (frequency 262 Hz) in its fundamental mode. (a) What should be the tension in the string? (b) What should be the mass of string S_2 so that it will produce A-sharp (frequency 466 Hz) as its fundamental? (c) To extend the range of your instrument, you include a fret located just under the strings but not normally touching them. How far from the upper end should you put this fret so that when you press S_1 tightly against it, this string will produce C-sharp (frequency 277 Hz) in its fundamental? That is, what is x in the figure? (d) If you press S_2 against the fret, what frequency of sound will it produce in its fundamental?

Figure **P15.52**



15.53 ••• CP A 5.00 m, 0.732 kg wire is used to support two uniform 235 N posts of equal length (**Fig. P15.53**). Assume that the wire is essentially horizontal and that the speed of sound is 344 m/s. A strong wind is blowing, causing the wire to vibrate in its 5th overtone. What are the frequency and wavelength of the sound this wire produces?

Figure **P15.53**



15.54 ••• CP You are exploring a newly discovered planet. The radius of the planet is 7.20×10^7 m. You suspend a lead weight from the lower end of a light string that is 4.00 m long and has mass 0.0280 kg. You measure that it takes 0.0685 s for a transverse pulse to travel from the lower end to the upper end of the string. On the earth, for the same string and lead weight, it takes 0.0390 s for a transverse pulse to travel the length of the string. The weight of the string is small enough that you ignore its effect on the tension in the string. Assuming that the mass of the planet is distributed with spherical symmetry, what is its mass?

15.55 •• For a string stretched between two supports, two successive standing-wave frequencies are 525 Hz and 630 Hz. There are other standing-wave frequencies lower than 525 Hz and higher than 630 Hz. If the speed of transverse waves on the string is 384 m/s, what is the length of the string? Assume that the mass of the wire is small enough for its effect on the tension in the wire to be ignored.

15.56 •• Transverse standing waves are produced on a string that has length 0.800 m and is held fixed at each end. Each standing-wave pattern has a node at the fixed ends plus additional nodes along the length of the string. You measure the frequencies f_n for standing waves that have n of these nodes along their length. The tension in the string is kept constant. You plot n versus f_n and find that your data lie close to a straight line that has slope 7.30×10^{-3} s. What is the speed of transverse waves on the string?

15.57 ••• CP A 1.80-m-long uniform bar that weighs 638 N is suspended in a horizontal position by two vertical wires that are attached to the ceiling. One wire is aluminum and the other is copper. The aluminum wire is attached to the left-hand end of the bar, and the copper wire is attached 0.40 m to the left of the right-hand end. Each wire has length 0.600 m and a circular cross section with radius 0.280 mm. What is the fundamental frequency of transverse standing waves for each wire?

15.58 •• CALC A transverse standing wave is set up on a string that is held fixed at both ends. The amplitude of the standing wave at an antinode is 1.80 mm and the speed of propagation of transverse waves on the string is 260 m/s. The string extends along the x-axis, with one of the fixed ends at $x = 0$, so that there is a node at $x = 0$. The smallest value of x where there is an antinode is $x = 0.150$ m. (a) What is the maximum transverse speed of a point on the string at an antinode? (b) What is the maximum transverse speed of a point on the string at $x = 0.075$ m?

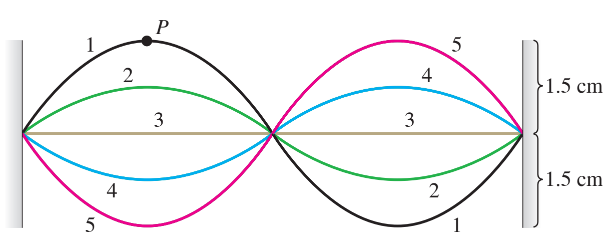
15.59 •• A horizontal wire is tied to supports at each end and vibrates in its second-overtone standing wave. The tension in the wire is 5.00 N, and the node-to-node distance in the standing wave is 6.28 cm. (a) What is the length of the wire? (b) A point at an antinode of the standing wave on the wire travels from its maximum upward displacement to its maximum downward displacement in 8.40 ms. What is the wire's mass?

15.60 ••• **CP** A vertical, 1.20 m length of 18 gauge (diameter of 1.024 mm) copper wire has a 100.0 N ball hanging from it. (a) What is the wavelength of the third harmonic for this wire? (b) A 500.0 N ball now *replaces* the original ball. What is the change in the wavelength of the third harmonic caused by replacing the light ball with the heavy one? (*Hint:* See Table 11.1 for Young's modulus.)

15.61 ••• A sinusoidal transverse wave travels on a string. The string has length 8.00 m and mass 6.00 g. The wave speed is 30.0 m/s, and the wavelength is 0.200 m. (a) If the wave is to have an average power of 50.0 W, what must be the amplitude of the wave? (b) For this same string, if the amplitude and wavelength are the same as in part (a), what is the average power for the wave if the tension is increased such that the wave speed is doubled?

15.62 •• A vibrating string 50.0 cm long is under a tension of 1.00 N. The results from five successive stroboscopic pictures are shown in **Fig. P15.62**. The strobe rate is set at 5000 flashes per minute, and observations reveal that the maximum displacement occurred at flashes 1 and 5 with no other maxima in between. (a) Find the period, frequency, and wavelength for the traveling waves on this string. (b) In what normal mode (harmonic) is the string vibrating? (c) What is the speed of the traveling waves on the string? (d) How fast is point P moving when the string is in (i) position 1 and (ii) position 3? (e) What is the mass of this string?

Figure **P15.62**



15.63 ••• **CP** A 1.005 m chain consists of small spherical beads, each with a mass of 1.00 g and a diameter of 5.00 mm, threaded on an elastic strand with negligible mass such that adjacent beads are separated by a center-to-center distance of 10.0 mm. There are beads at each end of the chain. The strand has a spring constant of 28.8 N/m. The chain is stretched horizontally on a frictionless tabletop to a length of 1.50 m, and the beads at both ends are fixed in place. (a) What is the linear mass density of the chain? (b) What is the tension in the chain? (c) With what speed would a pulse travel down the chain? (d) The chain is set vibrating and exhibits a standing-wave pattern with four antinodes. What is the frequency of this motion? (e) If the beads are numbered sequentially from 1 to 101, what are the numbers of the five beads that remain motionless? (f) The 13th bead has a maximum speed of 7.54 m/s. What is the amplitude of that bead's motion? (g) If $x_0 = 0$ corresponds to the center of the 1st bead and $x_{101} = 1.50$ m corresponds to the center of the 101st bead, what is the position x_n of the nth bead? (h) What is the maximum speed of the 30th bead?

15.64 •• A strong string of mass 3.00 g and length 2.20 m is tied to supports at each end and is vibrating in its fundamental mode. The maximum transverse speed of a point at the middle of the string is 9.00 m/s. The tension in the string is 330 N. (a) What is the amplitude of the standing wave at its antinode? (b) What is the magnitude of the maximum transverse acceleration of a point at the antinode?

15.65 •• A thin string 2.50 m in length is stretched with a tension of 90.0 N between two supports. When the string vibrates in its first overtone, a point at an antinode of the standing wave on the string has an amplitude of 3.50 cm and a maximum transverse speed of 28.0 m/s. (a) What is the string's mass? (b) What is the magnitude of the maximum transverse acceleration of this point on the string?

15.66 ••• **CALC** A guitar string is vibrating in its fundamental mode, with nodes at each end. The length of the segment of the string that is free to vibrate is 0.386 m. The maximum transverse acceleration of a point at the middle of the segment is 8.40×10^3 m/s^2 and the maximum transverse velocity is 3.80 m/s. (a) What is the amplitude of this standing wave? (b) What is the wave speed for the transverse traveling waves on this string?

15.67 ••• A uniform cylindrical steel wire, 55.0 cm long and 1.14 mm in diameter, is fixed at both ends. To what tension must it be adjusted so that, when vibrating in its first overtone, it produces the note D-sharp of frequency 311 Hz? Assume that it stretches an insignificant amount. (*Hint:* See Table 12.1.)

15.68 •• A string with both ends held fixed is vibrating in its third harmonic. The waves have a speed of 192 m/s and a frequency of 240 Hz. The amplitude of the standing wave at an antinode is 0.400 cm. (a) Calculate the amplitude at points on the string a distance of (i) 40.0 cm; (ii) 20.0 cm; and (iii) 10.0 cm from the left end of the string. (b) At each point in part (a), how much time does it take the string to go from its largest upward displacement to its largest downward displacement? (c) Calculate the maximum transverse velocity and the maximum transverse acceleration of the string at each of the points in part (a).

15.69 ••• **CP** A large rock that weighs 164.0 N is suspended from the lower end of a thin wire that is 3.00 m long. The density of the rock is 3200 kg/m^3. The mass of the wire is small enough that its effect on the tension in the wire can be ignored. The upper end of the wire is held fixed. When the rock is in air, the fundamental frequency for transverse standing waves on the wire is 42.0 Hz. When the rock is totally submerged in a liquid, with the top of the rock just below the surface, the fundamental frequency for the wire is 28.0 Hz. What is the density of the liquid?

15.70 •• (a) Estimate the tension you would need to apply to a standard small rubber band to stretch it between your fingers to a doubled length of 10 cm. (b) Such a rubber band has a typical mass of 0.10 g. Use this value to estimate the mass density of the stretched rubber band in SI units. (c) Use your estimated values to determine the expected frequency of vibration when the string is plucked. (d) Is this result realistic?

15.71 •• **Tuning an Instrument.** A musician tunes the C-string of her instrument to a fundamental frequency of 65.4 Hz. The vibrating portion of the string is 0.600 m long and has a mass of 14.4 g. (a) With what tension must the musician stretch it? (b) What percent increase in tension is needed to increase the frequency from 65.4 Hz to 73.4 Hz, corresponding to a rise in pitch from C to D?

15.72 • **Holding Up Under Stress.** A string or rope will break apart if it is placed under too much tensile stress [see Eq. (11.8)]. Thicker ropes can withstand more tension without breaking because the thicker the rope, the greater the cross-sectional area and the smaller the stress. One type of steel has density 7800 kg/m^3 and will break if the tensile stress exceeds 7.0×10^8 N/m^2. You want to make a guitar string from 4.0 g of this type of steel. In use, the guitar string must be able to withstand a tension of 900 N without breaking. Your job is to determine (a) the maximum length and minimum radius the string can have; (b) the highest possible fundamental frequency of standing waves on this string, if the entire length of the string is free to vibrate.

15.73 •• **DATA** In your physics lab, an oscillator is attached to one end of a horizontal string. The other end of the string passes over a frictionless pulley. You suspend a mass M from the free end of the string, producing tension Mg in the string. The oscillator produces transverse waves of frequency f on the string. You don't vary this frequency during the experiment, but you try strings with three different linear mass densities μ. You also keep a fixed distance between the end of the string where the oscillator is attached and the point where the string is in contact with the pulley's rim. To produce standing waves on the string, you vary M; then you measure the node-to-node distance d for each standing-wave pattern and obtain the following data:

String	A	A	B	B	C
μ **(g/cm)**	0.0260	0.0260	0.0374	0.0374	0.0482
M **(g)**	559	249	365	207	262
d **(cm)**	48.1	31.9	32.0	24.2	23.8

(a) Explain why you obtain only certain values of d. (b) Graph μd^2 (in kg $\cdot$ m) versus M (in kg). Explain why the data plotted this way should fall close to a straight line. (c) Use the slope of the best straight-line fit to the data to determine the frequency f of the waves produced on the string by the oscillator. Take $g = 9.80 \text{ m/s}^2$. (d) For string A ($\mu = 0.0260$ g/cm), what value of M (in grams) would be required to produce a standing wave with a node-to-node distance of 24.0 cm? Use the value of f that you calculated in part (c).

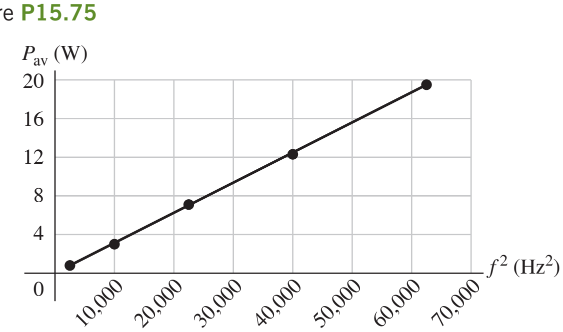
15.74 •• DATA *Scale length* is the length of the part of a guitar string that is free to vibrate. A standard value of scale length for an acoustic guitar is 25.5 in. The frequency of the fundamental standing wave on a string is determined by the string's scale length, tension, and linear mass density. The standard frequencies f to which the strings of a six-string guitar are tuned are given in the table:

String	E2	A2	D3	G3	B3	E4
f **(Hz)**	82.4	110.0	146.8	196.0	246.9	329.6

Assume that a typical value of the tension of a guitar string is 78.0 N (although tension varies somewhat for different strings). (a) Calculate the linear mass density μ (in g/cm) for the E2, G3, and E4 strings. (b) Just before your band is going to perform, your G3 string breaks. The only replacement string you have is an E2. If your strings have the linear mass densities calculated in part (a), what must be the tension in the replacement string to bring its fundamental frequency to the G3 value of 196.0 Hz?

15.75 •• DATA You are measuring the frequency dependence of the average power P_{av} transmitted by traveling waves on a wire. In your experiment you use a wire with linear mass density 3.5 g/m. For a transverse wave on the wire with amplitude 4.0 mm, you measure P_{av} (in watts) as a function of the frequency f of the wave (in Hz). You have chosen to plot P_{av} as a function of f^2 (**Fig. P15.75**). (a) Explain why values of P_{av} plotted versus f^2 should be well fit by a straight line. (b) Use the slope of the straight-line fit to the data shown in Fig. P15.75 to calculate the speed of the waves. (c) What angular frequency ω would result in $P_{av} = 10.0$ W?

Figure **P15.75**

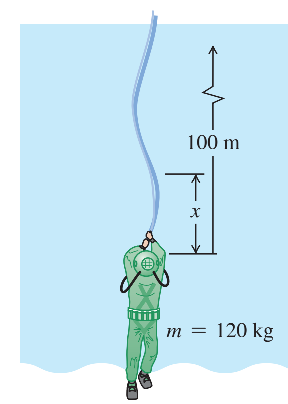


CHALLENGE PROBLEMS

15.76 •• A rectangular neoprene sheet has width $W = 1.00$ m and length $L = 4.00$ m. The two shorter edges are affixed to rigid steel bars that are used to stretch the sheet taut and horizontal. The force applied to either end of the sheet is $F = 81.0$ N. The sheet has a total mass $M = 4.00$ kg. The left edge of the sheet is wiggled vertically in a uniform sinusoidal motion with amplitude $A = 10.0$ cm and frequency $f = 1.00$ Hz. This sends waves spanning the width of the sheet rippling from left to right. The right side of the sheet moves upward and downward freely as these waves complete their traversal. (a) Use a two-dimensional generalization of the discussion in Section 15.4 to derive an expression for the velocity with which the waves move along the sheet in terms of generic values of $W, L, F, M, f,$ and A. What is the value of this speed for the specified choices of these parameters? (b) If the positive x-axis is oriented rightward and the steel bars are parallel to the y-axis, the height of the sheet may be characterized as $z(x, y) = A \sin(kx - \omega t)$. What is the value of the wave number k? (c) Write down an expression with generic parameters for the rate of rightward energy transfer by the slice of sheet at a given value of x at generic time t. (d) The power at $x = 0$ is supplied by the agent wiggling the left bar upward and downward. How much energy is supplied each second by that agent? Express your answer in terms of generic parameters and also as a specific energy for the given parameters.

15.77 ••• CP CALC A deep-sea diver is suspended beneath the surface of Loch Ness by a 100-m-long cable that is attached to a boat on the surface (**Fig. P15.77**). The diver and his suit have a total mass of 120 kg and a volume of 0.0800 m^3. The cable has a diameter of 2.00 cm and a linear mass density of $\mu = 1.10$ kg/m. The diver thinks he sees something moving in the murky depths and jerks the end of the cable back and forth to send transverse waves up the cable as a signal to his companions in the boat. (a) What is the tension in the cable at its lower end, where it is attached to the diver? Do not forget to include the buoyant force that the water (density 1000 kg/m^3) exerts on him. (b) Calculate the tension in the cable a distance x above the diver. In your calculation, include the buoyant force on the cable. (c) The speed of transverse waves on the cable is given by $v = \sqrt{F/\mu}$ [Eq. (15.14)]. The speed therefore varies along the cable, since the tension is not constant. (This expression ignores the damping force that the water exerts on the moving cable.) Integrate to find the time required for the first signal to reach the surface.

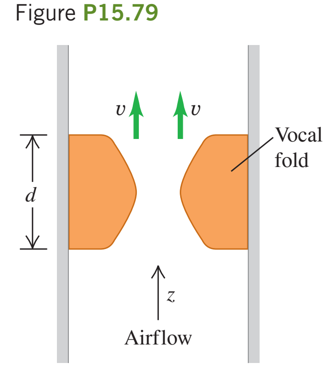
Figure **P15.77**



MCAT-STYLE PASSAGE PROBLEMS

BIO Waves on Vocal Folds. In the larynx, sound is produced by the vibration of the *vocal folds* (also called "vocal cords"). The accompanying figure is a cross section of the vocal tract at one instant in time. Air flows upward (in the $+z$-direction) through the vocal tract, causing a transverse wave to propagate vertically upward along the surface of the vocal folds. In a typical adult male, the thickness of the vocal folds in the direction of airflow is $d = 2.0$ mm. High-speed photography shows that for a frequency of vibration of $f = 125$ Hz, the wave along the surface of the vocal folds travels upward at a speed of $v = 375$ cm/s. Use t for time, z for displacement in the $+z$-direction, and λ for wavelength.

Figure **P15.79**

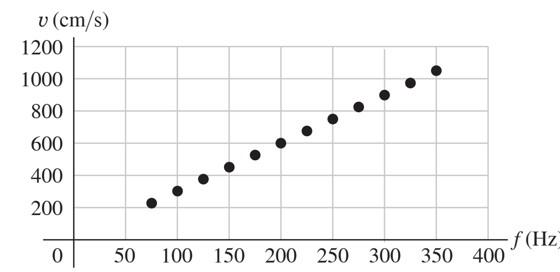


15.78 What is the wavelength of the wave that travels on the surface of the vocal folds when they are vibrating at frequency f? (a) 2.0 mm; (b) 3.3 mm; (c) 0.50 cm; (d) 3.0 cm.

15.79 Which of these is a possible mathematical description of the wave in Problem 15.78? (a) $A\sin[2\pi f(t + z/v)]$; (b) $A\sin[2\pi f(t - z/v)]$; (c) $A\sin(2\pi ft)\cos(2\pi z/\lambda)$; (d) $A\sin(2\pi ft)\sin(2\pi z/\lambda)$.

15.80 The wave speed is measured for different vibration frequencies. A graph of the wave speed as a function of frequency (**Fig. P15.80**) indicates that as the frequency increases, the wavelength (a) increases; (b) decreases; (c) doesn't change; (d) becomes undefined.

Figure **P15.80**



ANSWERS

Chapter Opening Question ?

(iii) The power of a mechanical wave depends on both its amplitude and its frequency [see Eq. (15.25)].

Key Example VARIATION Problems

VP15.3.1 (a) 0.196 Hz (b) 5.98 m/s (c) 1.23 rad/s (d) 0.206 m^{-1}
VP15.3.2 (a) $T = 8.00 \times 10^{-3}$ s, $\lambda = 1.96$ m (b) $f = 81.7$ Hz, $\omega = 513$ rad/s
VP15.3.3 (a) 111 m/s (b) 4.44 m (c) $y(x = +2.50 \text{ m}, t) = (5.00 \text{ mm})\cos[(3.54 \text{ rad}) - (157 \text{ rad/s})t]$
VP15.3.4 (a) 36.0 m/s (b) 0.0193 kg/m
VP15.5.1 (a) 10.9 rad/s (b) 0.294 W
VP15.5.2 2.50×10^{-3} m
VP15.5.3 (a) 0.159 W/m^2 (b) 3.76 m
VP15.5.4 6.94×10^{-12} J
VP15.8.1 (a) 64.0 m/s (b) 2.25 m/s (c) 3.62×10^3 m/s^2
VP15.8.2 (a) 251 m/s (b) 145 N
VP15.8.3 (a) 1.20 m (b) 80.0 Hz (c) 1.90×10^{-2} kg/m
VP15.8.4 (a) 76.7 N (b) 1.34 m

Bridging Problem

(a) $F(r) = \dfrac{m\omega^2}{2L}(L^2 - r^2)$

(b) $v(r) = \omega\sqrt{\dfrac{L^2 - r^2}{2}}$

(c) $\dfrac{\pi}{\omega\sqrt{2}}$

? The sound from a horn travels more slowly on a cold winter day high in the mountains than on a warm summer day at sea level. This is because at high elevations in winter, the air has lower (i) pressure; (ii) density; (iii) humidity; (iv) temperature; (v) mass per mole.

16 Sound and Hearing

Of all the mechanical waves that occur in nature, the most important in our everyday lives are longitudinal waves in a medium—usually air—called *sound* waves. The reason is that the human ear is tremendously sensitive and can detect sound waves even of very low intensity. The ability to hear an unseen nocturnal predator was essential to the survival of our ancestors, so it is no exaggeration to say that we humans owe our existence to our highly evolved sense of hearing.

In Chapter 15 we described mechanical waves primarily in terms of displacement; however, because the ear is primarily sensitive to changes in pressure, it's often more appropriate to describe sound waves in terms of *pressure* fluctuations. We'll study the relationships among displacement, pressure fluctuation, and intensity and the connections between these quantities and human sound perception.

When a source of sound or a listener moves through the air, the listener may hear a frequency different from the one emitted by the source. This is the Doppler effect, which has important applications in medicine and technology.

LEARNING OUTCOMES

In this chapter, you'll learn...

- **16.1** How to describe a sound wave in terms of either particle displacements or pressure fluctuations.
- **16.2** How to calculate the speed of sound waves in different materials.
- **16.3** How to calculate the intensity of a sound wave.
- **16.4** What determines the particular frequencies of sound produced by an organ or a flute.
- **16.5** How resonance occurs in musical instruments.
- **16.6** What happens when sound waves from different sources overlap.
- **16.7** How to describe what happens when two sound waves of slightly different frequencies are combined.
- **16.8** Why the pitch of a siren changes as it moves past you.
- **16.9** Why an airplane flying faster than sound produces a shock wave.

You'll need to review...

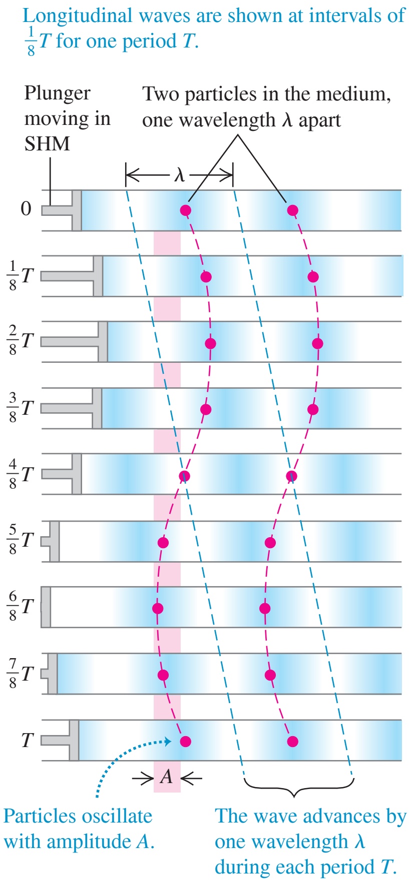
- **6.4** Power.
- **8.1** The impulse–momentum theorem.
- **11.4** Bulk modulus and Young's modulus.
- **12.2** Gauge pressure and absolute pressure.
- **14.8** Forced oscillations and resonance.
- **15.1–15.8** Mechanical waves.

16.1 SOUND WAVES

The most general definition of **sound** is a longitudinal wave in a medium. Our main concern is with sound waves in air, but sound can travel through any gas, liquid, or solid. You may be all too familiar with the propagation of sound through a solid if your neighbor's stereo speakers are right next to your wall.

The simplest sound waves are sinusoidal waves, which have definite frequency, amplitude, and wavelength. The human ear is sensitive to waves in the frequency range from about 20 to 20,000 Hz, called the **audible range,** but we also use the term "sound" for similar waves with frequencies above **(ultrasonic)** and below **(infrasonic)** the range of human hearing.

Figure **16.1** A sinusoidal longitudinal wave traveling to the right in a fluid. (Compare to Fig. 15.7.)



Sound waves usually travel outward in all directions from the source of sound, with an amplitude that depends on the direction and distance from the source. We'll return to this point in the next section. For now, we concentrate on the idealized case of a sound wave that propagates in the positive x-direction only. As we discussed in Section 15.3, for such a wave, the wave function $y(x, t)$ gives the instantaneous displacement y of a particle in the medium at position x at time t. If the wave is sinusoidal, we can express it by using Eq. (15.7):

$$y(x, t) = A\cos(kx - \omega t) \qquad \text{(sound wave propagating in the } +x \text{ direction)} \tag{16.1}$$

In a longitudinal wave the displacements are *parallel* to the direction of travel of the wave, so distances x and y are measured parallel to each other, not perpendicular as in a transverse wave. The amplitude A is the maximum displacement of a particle in the medium from its equilibrium position (**Fig. 16.1**). Hence A is also called the **displacement amplitude.**

Sound Waves as Pressure Fluctuations

We can also describe sound waves in terms of variations of *pressure* at various points. In a sinusoidal sound wave in air, the pressure fluctuates sinusoidally above and below atmospheric pressure p_a with the same frequency as the motions of the air particles. The human ear operates by sensing such pressure variations. A sound wave entering the ear canal exerts a fluctuating pressure on one side of the eardrum; the air on the other side of the eardrum, vented to the outside by the Eustachian tube, is at atmospheric pressure. The pressure difference on the two sides of the eardrum sets it into motion. Microphones and similar devices also usually sense pressure differences, not displacements.

Let $p(x, t)$ be the instantaneous pressure fluctuation in a sound wave at any point x at time t. That is, $p(x, t)$ is the amount by which the pressure *differs* from normal atmospheric pressure p_a. Think of $p(x, t)$ as the *gauge pressure* defined in Section 12.2; it can be either positive or negative. The *absolute* pressure at a point is then $p_a + p(x, t)$.

To see the connection between the pressure fluctuation $p(x, t)$ and the displacement $y(x, t)$ in a sound wave propagating in the $+x$-direction, consider an imaginary cylinder of a wave medium (gas, liquid, or solid) with cross-sectional area S and its axis along the direction of propagation (**Fig. 16.2**). When no sound wave is present, the cylinder has length Δx and volume $V = S\,\Delta x$, as shown by the shaded volume in Fig. 16.2. When a wave is present, at time t the end of the cylinder that is initially at x is displaced by $y_1 = y(x, t)$, and the end that is initially at $x + \Delta x$ is displaced by $y_2 = y(x + \Delta x, t)$; this is shown by the red lines. If $y_2 > y_1$, as shown in Fig. 16.2, the cylinder's volume has increased, which causes a decrease in pressure. If $y_2 < y_1$, the cylinder's volume has decreased and the pressure has increased. If $y_2 = y_1$, the cylinder is simply shifted to the left or right; there is no volume change and no pressure fluctuation. The pressure fluctuation depends on the *difference* between the displacements at neighboring points in the medium.

Quantitatively, the change in volume ΔV of the cylinder is

$$\Delta V = S(y_2 - y_1) = S[y(x + \Delta x, t) - y(x, t)]$$

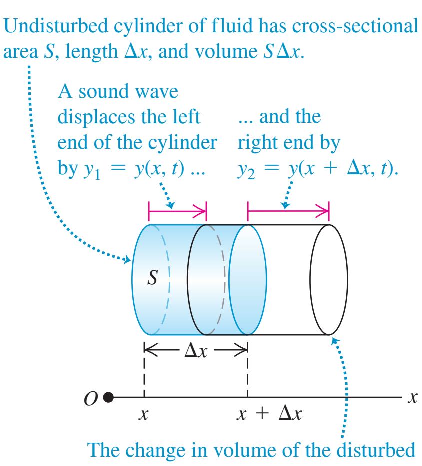
In the limit as $\Delta x \to 0$, the fractional change in volume dV/V (volume change divided by original volume) is

$$\frac{dV}{V} = \lim_{\Delta x \to 0} \frac{S[y(x + \Delta x, t) - y(x, t)]}{S\,\Delta x} = \frac{\partial y(x, t)}{\partial x} \tag{16.2}$$

The fractional volume change is related to the pressure fluctuation by the bulk modulus B, which by definition [Eq. (11.3)] is $B = -p(x, t)/(dV/V)$ (see Section 11.4). Solving for $p(x, t)$, we have

$$p(x, t) = -B\frac{\partial y(x, t)}{\partial x} \tag{16.3}$$

Figure **16.2** As a sound wave propagates along the x-axis, the left and right ends undergo different displacements y_1 and y_2.



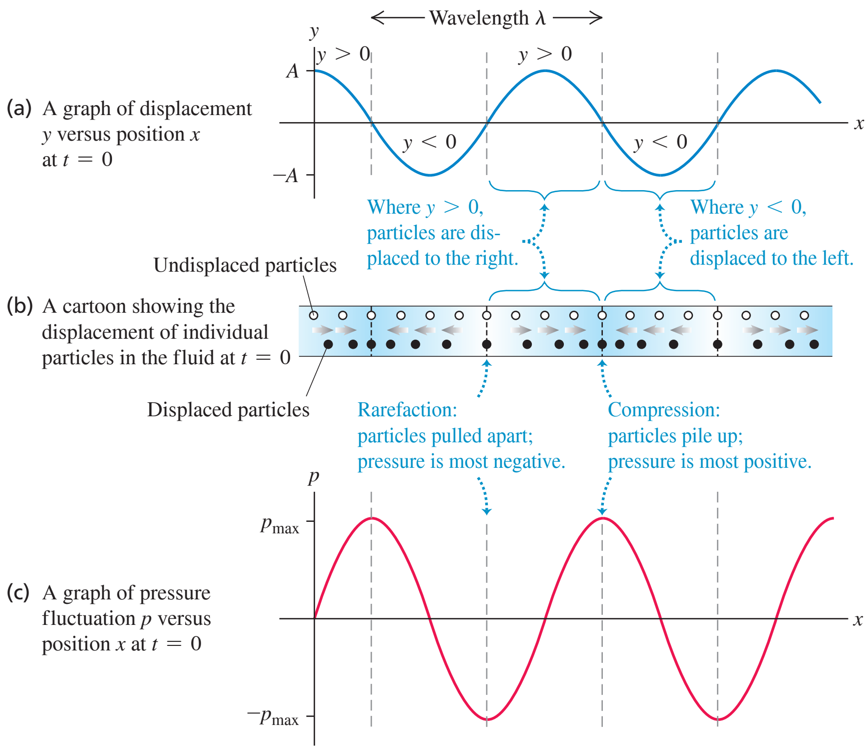


Figure **16.3** Three ways to describe a sound wave.

The negative sign arises because when $\partial y(x, t)/\partial x$ is positive, the displacement is greater at $x + \Delta x$ than at x, corresponding to an increase in volume, a decrease in pressure, and a negative pressure fluctuation.

When we evaluate $\partial y(x, t)/\partial x$ for the sinusoidal wave of Eq. (16.1), we find

$$p(x, t) = BkA \sin(kx - \omega t) \tag{16.4}$$

Figure 16.3 shows $y(x, t)$ and $p(x, t)$ for a sinusoidal sound wave at $t = 0$. It also shows how individual particles of the wave are displaced at this time. While $y(x, t)$ and $p(x, t)$ describe the same wave, these two functions are one-quarter cycle out of phase: At any time, the displacement is greatest where the pressure fluctuation is zero, and vice versa. In particular, note that the compressions (points of greatest pressure and density) and rarefactions (points of lowest pressure and density) are points of *zero* displacement.

Equation (16.4) shows that the quantity BkA represents the maximum pressure fluctuation. We call this the **pressure amplitude,** denoted by p_{max}:

Bulk modulus of medium

Pressure amplitude, sinusoidal sound wave

$$p_{max} = BkA \tag{16.5}$$

Displacement amplitude

Wave number $= 2\pi/\lambda$

Waves of shorter wavelength λ (larger wave number $k = 2\pi/\lambda$) have greater pressure variations for a given displacement amplitude because the maxima and minima are squeezed closer together. A medium with a large value of bulk modulus B is less compressible and so requires a greater pressure amplitude for a given volume change (that is, a given displacement amplitude).

CAUTION **Graphs of a sound wave** The graphs in Fig. 16.3 show the wave at only *one* instant of time. Because the wave is propagating in the $+x$-direction, as time goes by the wave patterns described by the functions $y(x, t)$ and $p(x, t)$ move to the right at the wave speed $v = \omega/k$. The particles, by contrast, simply oscillate back and forth in simple harmonic motion as shown in Fig. 16.1.

EXAMPLE 16.1 Amplitude of a sound wave in air

In a sinusoidal sound wave of moderate loudness, the maximum pressure variations are about 3.0×10^{-2} Pa above and below atmospheric pressure. Find the corresponding maximum displacement if the frequency is 1000 Hz. In air at normal atmospheric pressure and density, the speed of sound is 344 m/s and the bulk modulus is 1.42×10^5 Pa.

IDENTIFY and SET UP This problem involves the relationship between two ways of describing a sound wave: in terms of displacement and in terms of pressure. The target variable is the displacement amplitude A. We are given the pressure amplitude p_{max}, wave speed v, frequency f, and bulk modulus B. Equation (16.5) relates the target variable A to p_{max}. We use $\omega = vk$ [Eq. (15.6)] to determine the wave number k from v and the angular frequency $\omega = 2\pi f$.

EXECUTE From Eq. (15.6),

$$k = \frac{\omega}{v} = \frac{2\pi f}{v} = \frac{(2\pi \text{ rad})(1000 \text{ Hz})}{344 \text{ m/s}} = 18.3 \text{ rad/m}$$

Then from Eq. (16.5), the maximum displacement is

$$A = \frac{p_{max}}{Bk} = \frac{3.0 \times 10^{-2} \text{ Pa}}{(1.42 \times 10^5 \text{ Pa})(18.3 \text{ rad/m})} = 1.2 \times 10^{-8} \text{ m}$$

EVALUATE This displacement amplitude is only about $\frac{1}{100}$ the size of a human cell. The ear actually senses pressure fluctuations; it detects these minuscule displacements only indirectly.

KEYCONCEPT In a sound wave, the pressure amplitude (maximum pressure fluctuation) and displacement amplitude (maximum displacement of a particle in the medium) are proportional to each other. The proportionality constant depends on the wavelength of the sound and the bulk modulus of the medium.

EXAMPLE 16.2 Amplitude of a sound wave in the inner ear

A sound wave that enters the human ear sets the eardrum into oscillation, which in turn causes oscillation of the *ossicles,* a chain of three tiny bones in the middle ear (**Fig. 16.4**). The ossicles transmit this oscillation to the fluid (mostly water) in the inner ear; there the fluid motion disturbs hair cells that send nerve impulses to the brain with information about the sound. The area of the moving part of the eardrum is about 43 mm^2, and that of the stapes (the smallest of the ossicles) where it connects to the inner ear is about 3.2 mm^2. For the sound in Example 16.1, determine (a) the pressure amplitude and (b) the displacement amplitude of the wave in the fluid of the inner ear, in which the speed of sound is 1500 m/s.

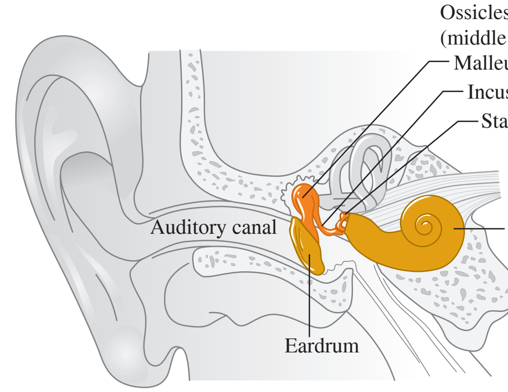
IDENTIFY and SET UP Although the sound wave here travels in liquid rather than air, the same principles and relationships among the properties of the wave apply. We can ignore the mass of the tiny ossicles (about 58 mg $= 5.8 \times 10^{-5}$ kg), so the force they exert on the inner-ear fluid is the same as that exerted on the eardrum and ossicles by the incident sound wave. (In Chapters 4 and 5 we used the same idea to say that the tension is the same at either end of a massless rope.) Hence the pressure amplitude in the inner ear, $p_{max\,(inner\ ear)}$, is greater than in the outside air, $p_{max\,(air)}$, because the same force is exerted on a smaller area (the area of the stapes versus the area of the eardrum). Given $p_{max\,(inner\ ear)}$, we find the displacement amplitude $A_{inner\ ear}$ from Eq. (16.5).

EXECUTE (a) From the area of the eardrum and the pressure amplitude in air found in Example 16.1, the maximum force exerted by the sound wave in air on the eardrum is $F_{max} = p_{max\,(air)}S_{eardrum}$. Then

$$p_{max\,(inner\ ear)} = \frac{F_{max}}{S_{stapes}} = p_{max\,(air)}\frac{S_{eardrum}}{S_{stapes}}$$

$$= (3.0 \times 10^{-2} \text{ Pa})\frac{43 \text{ mm}^2}{3.2 \text{ mm}^2} = 0.40 \text{ Pa}$$

Figure 16.4 The anatomy of the human ear. The middle ear is the size of a small marble; the ossicles (incus, malleus, and stapes) are the smallest bones in the human body.



(b) To find the maximum displacement $A_{inner\ ear}$, we use $A = p_{max}/Bk$ as in Example 16.1. The inner-ear fluid is mostly water, which has a much greater bulk modulus B than air. From Table 11.2 the compressibility of water (unfortunately also called k) is 45.8×10^{-11} Pa^{-1}, so $B_{fluid} = 1/(45.8 \times 10^{-11} \text{ Pa}^{-1}) = 2.18 \times 10^9$ Pa.

The wave in the inner ear has the same angular frequency ω as the wave in the air because the air, eardrum, ossicles, and inner-ear fluid all oscillate together (see Example 15.8 in Section 15.8). But because the wave speed v is greater in the inner ear than in the air (1500 m/s versus 344 m/s), the wave number $k = \omega/v$ is smaller. Using the value of ω from Example 16.1, we find

$$k_{inner\ ear} = \frac{\omega}{v_{inner\ ear}} = \frac{(2\pi \text{ rad})(1000 \text{ Hz})}{1500 \text{ m/s}} = 4.2 \text{ rad/m}$$

Putting everything together, we have

$$A_{\text{inner ear}} = \frac{p_{\max\,(\text{inner ear})}}{B_{\text{fluid}}k_{\text{inner ear}}} = \frac{0.40\ \text{Pa}}{(2.18 \times 10^9\ \text{Pa})(4.2\ \text{rad/m})}$$

$$= 4.4 \times 10^{-11}\ \text{m}$$

EVALUATE In part (a) we see that the ossicles increase the pressure amplitude by a factor of $(43\ \text{mm}^2)/(3.2\ \text{mm}^2) = 13$. This amplification helps give the human ear its great sensitivity.

The displacement amplitude in the inner ear is even smaller than in the air. But *pressure* variations within the inner-ear fluid are what set the hair cells into motion, so what matters is that the pressure amplitude is larger in the inner ear than in the air.

KEYCONCEPT When a sound wave travels from one medium into a different medium, the wave frequency and angular frequency remain the same. The wave number and wavelength can change, however, as can the pressure amplitude and displacement amplitude.

Perception of Sound Waves

The physical characteristics of a sound wave are directly related to the perception of that sound by a listener. For a given frequency, the greater the pressure amplitude of a sinusoidal sound wave, the greater the perceived **loudness.** The relationship between pressure amplitude and loudness is not a simple one, and it varies from one person to another. One important factor is that the ear is not equally sensitive to all frequencies in the audible range. A sound at one frequency may seem louder than one of equal pressure amplitude at a different frequency. At 1000 Hz the minimum pressure amplitude that can be perceived with normal hearing is about 3×10^{-5} Pa; to produce the same loudness at 200 Hz or 15,000 Hz requires about 3×10^{-4} Pa. Perceived loudness also depends on the health of the ear. Age usually brings a loss of sensitivity at high frequencies.

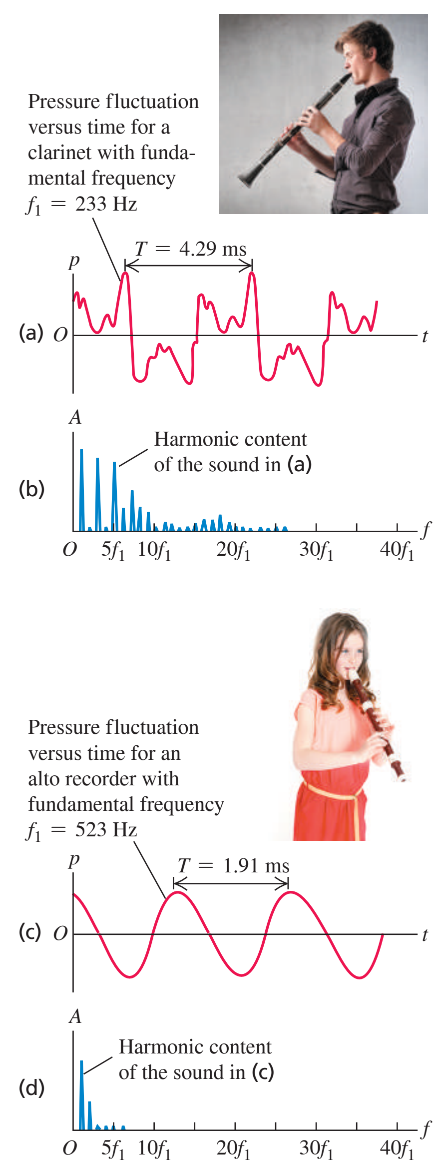
The frequency of a sound wave is the primary factor in determining the **pitch** of a sound, the quality that lets us classify the sound as "high" or "low." The higher the frequency of a sound (within the audible range), the higher the pitch that a listener will perceive. Pressure amplitude also plays a role in determining pitch. When a listener compares two sinusoidal sound waves with the same frequency but different pressure amplitudes, the one with the greater pressure amplitude is usually perceived as louder but also as slightly lower in pitch.

Musical sounds have wave functions that are more complicated than a simple sine function. **Figure 16.5a** shows the pressure fluctuation in the sound wave produced by a clarinet. The pattern is so complex because the column of air in a wind instrument like a clarinet vibrates at a fundamental frequency and at many harmonics at the same time. (In Section 15.8, we described this same behavior for a string that has been plucked, bowed, or struck. We'll examine the physics of wind instruments in Section 16.4.) The sound wave produced in the surrounding air has a similar amount of each harmonic—that is, a similar *harmonic content.* Figure 16.5b shows the harmonic content of the sound of a clarinet. The mathematical process of translating a pressure-time graph like Fig. 16.5a into a graph of harmonic content like Fig. 16.5b is called *Fourier analysis.*

Two tones produced by different instruments might have the same fundamental frequency (and thus the same pitch) but sound different because of different harmonic content. The difference in sound is called **timbre** and is often described in subjective terms such as reedy, mellow, and tinny. A tone that is rich in harmonics, like the clarinet tone in Figs. 16.5a and 16.5b, usually sounds thin and "reedy," while a tone containing mostly a fundamental, like the alto recorder tone in Figs. 16.5c and 16.5d, is more mellow and flute-like. The same principle applies to the human voice, which is another wind instrument; the vowels "a" and "e" sound different because of differences in harmonic content.

Another factor in determining tone quality is the behavior at the beginning (*attack*) and end (*decay*) of a tone. A piano tone begins with a thump and then dies away gradually. A harpsichord tone, in addition to having different harmonic content, begins much more quickly with a click, and the higher harmonics begin before the lower ones. When the key is released, the sound also dies away much more rapidly with a harpsichord than with a piano. Similar effects are present in other musical instruments.

Figure **16.5** Different representations of the sound of (a), (b) a clarinet and (c), (d) an alto recorder. (Graphs adapted from R.E. Berg and D.G. Stork, *The Physics of Sound,* Prentice-Hall, 1982.)



BIO APPLICATION Hearing Loss from Amplified Sound Due to exposure to highly amplified music, many young musicians have suffered permanent ear damage and have hearing typical of persons 65 years of age. Headphones for personal music players used at high volume pose similar threats to hearing. Be careful!

Unlike the tones made by musical instruments, **noise** is a combination of *all* frequencies, not just frequencies that are integer multiples of a fundamental frequency. (An extreme case is "white noise," which contains equal amounts of all frequencies across the audible range.) Examples include the sound of the wind and the hissing sound you make in saying the consonant "s."

TEST YOUR UNDERSTANDING OF SECTION 16.1 You use an electronic signal generator to produce a sinusoidal sound wave in air. You then increase the frequency of the wave from 100 Hz to 400 Hz while keeping the pressure amplitude constant. What effect does this have on the displacement amplitude of the sound wave? (i) It becomes four times greater; (ii) it becomes twice as great; (iii) it is unchanged; (iv) it becomes $\frac{1}{2}$ as great; (v) it becomes $\frac{1}{4}$ as great.

ANSWER

❚ **(v)** From Eq. (16.5), the displacement amplitude is $A = p_{max}/Bk$. The pressure amplitude p_{max} and bulk modulus B remain the same, but the frequency f increases by a factor of 4. Hence the wave number $k = \omega/v = 2\pi f/v$ also increases by a factor of 4. Since A is inversely proportional to k, the displacement amplitude becomes $\frac{1}{4}$ as great. In other words, at higher frequency a smaller maximum displacement is required to produce the same maximum pressure fluctuation.

16.2 SPEED OF SOUND WAVES

We found in Section 15.4 that the speed v of a transverse wave on a string depends on the string tension F and the linear mass density μ: $v = \sqrt{F/\mu}$. What, we may ask, is the corresponding expression for the speed of sound waves in a gas or liquid? On what properties of the medium does the speed depend?

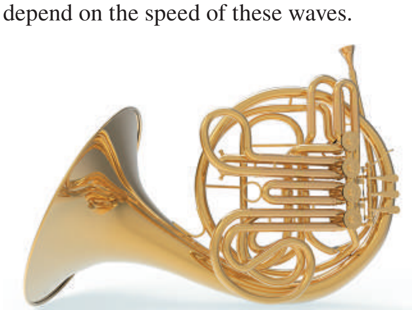
We can make an educated guess about these questions by remembering a claim that we made in Section 15.4: For mechanical waves in general, the expression for the wave speed is of the form

$$v = \sqrt{\frac{\text{Restoring force returning the system to equilibrium}}{\text{Inertia resisting the return to equilibrium}}}$$

A sound wave in a bulk fluid causes compressions and rarefactions of the fluid, so the restoring-force term in the above expression must be related to how difficult it is to compress the fluid. This is precisely what the bulk modulus B of the medium tells us. According to Newton's second law, inertia is related to mass. The "massiveness" of a bulk fluid is described by its density, or mass per unit volume, ρ. Hence we expect that the speed of sound waves should be of the form $v = \sqrt{B/\rho}$.

To check our guess, we'll derive the speed of sound waves in a fluid in a pipe. This is a situation of some importance, since all musical wind instruments are pipes in which a longitudinal wave (sound) propagates in a fluid (air) (**Fig. 16.6**). Human speech works on the same principle; sound waves propagate in your vocal tract, which is an air-filled pipe connected to the lungs at one end (your larynx) and to the outside air at the other end (your mouth). The steps in our derivation are completely parallel to those we used in Section 15.4 to find the speed of transverse waves.

Figure **16.6** When a wind instrument like this French horn is played, sound waves propagate through the air within the instrument's pipes. The properties of the sound that emerges from the large bell depend on the speed of these waves.



Speed of Sound in a Fluid

Figure 16.7 shows a fluid with density ρ in a pipe with cross-sectional area A. In equilibrium (Fig. 16.7a), the fluid is at rest and under a uniform pressure p. We take the x-axis along the length of the pipe. This is also the direction in which we make a longitudinal wave propagate, so the displacement y is also measured along the pipe, just as in Section 16.1 (see Fig. 16.2).

At time $t = 0$ we start the piston at the left end moving toward the right with constant speed v_y. This initiates a wave motion that travels to the right along the length of the pipe, in which successive sections of fluid begin to move and become compressed at successively later times.

Figure 16.7b shows the fluid at time t. All portions of fluid to the left of point P are moving to the right with speed v_y, and all portions to the right of P are still at rest. The boundary between the moving and stationary portions travels to the right with a speed equal to the speed of propagation or wave speed v. At time t the piston has moved a distance $v_y t$, and the boundary has advanced a distance vt. As with a transverse disturbance in a string, we can compute the speed of propagation from the impulse–momentum theorem.

The quantity of fluid set in motion in time t originally occupied a section of the cylinder with length vt, cross-sectional area A, volume vtA, and mass ρvtA. Its longitudinal momentum (that is, momentum along the length of the pipe) is

$$\text{Longitudinal momentum} = (\rho vtA)v_y$$

Next we compute the increase of pressure, Δp, in the moving fluid. The original volume of the moving fluid, Avt, has decreased by an amount $Av_y t$. From the definition of the bulk modulus B, Eq. (11.13) in Section 11.5,

$$B = \frac{-\text{Pressure change}}{\text{Fractional volume change}} = \frac{-\Delta p}{-Av_y t/Avt} \qquad \text{and} \qquad \Delta p = B\frac{v_y}{v}$$

The pressure in the moving fluid is $p + \Delta p$, and the force exerted on it by the piston is $(p + \Delta p)A$. The net force on the moving fluid (see Fig. 16.7b) is ΔpA, and the longitudinal impulse is

$$\text{Longitudinal impulse} = \Delta pAt = B\frac{v_y}{v}At$$

Because the fluid was at rest at time $t = 0$, the change in momentum up to time t is equal to the momentum at that time. Applying the impulse–momentum theorem (see Section 8.1), we find

$$B\frac{v_y}{v}At = \rho vtAv_y \tag{16.6}$$

When we solve this expression for v, we get

Speed of a longitudinal wave in a fluid
$$v = \sqrt{\frac{B}{\rho}} \tag{16.7}$$
B: Bulk modulus of fluid
ρ: Density of fluid

which agrees with our educated guess.

While we derived Eq. (16.7) for waves in a pipe, it also applies to longitudinal waves in a bulk fluid, including sound waves traveling in air or water.

Speed of Sound in a Solid

When a longitudinal wave propagates in a *solid* rod or bar, the situation is somewhat different. The rod expands sideways slightly when it is compressed longitudinally, while a fluid in a pipe with constant cross section cannot move sideways. Using the same kind of reasoning that led us to Eq. (16.7), we can show that the speed of a longitudinal pulse in the rod is given by

Speed of a longitudinal wave in a solid rod
$$v = \sqrt{\frac{Y}{\rho}} \tag{16.8}$$
Y: Young's modulus of rod material
ρ: Density of rod material

We defined Young's modulus in Section 11.4.

CAUTION **Solid rods vs. bulk solids** Equation (16.8) applies to only rods whose sides are free to bulge and shrink a little as the wave travels. It does not apply to longitudinal waves in a *bulk* solid because sideways motion in any element of material is prevented by the surrounding material. The speed of longitudinal waves in a bulk solid depends on the density, the bulk modulus, and the *shear* modulus.

Note that Eqs. (16.7) and (16.8) are valid for sinusoidal and other periodic waves, not just for the special case discussed here.

Table 16.1 lists the speed of sound in several bulk materials. Sound waves travel more slowly in lead than in aluminum or steel because lead has a lower bulk modulus and shear modulus and a higher density.

Figure 16.7 A sound wave propagating in a fluid confined to a tube. (a) Fluid in equilibrium. (b) A time t after the piston begins moving to the right at speed v_y, the fluid between the piston and point P is in motion. The speed of sound waves is v.

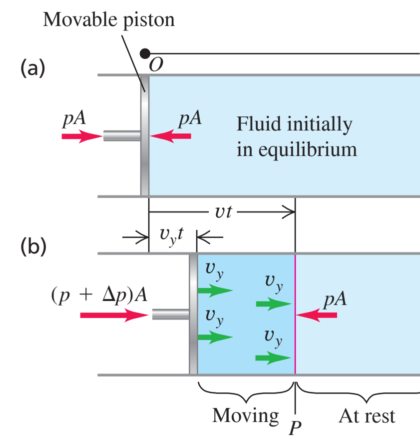


TABLE 16.1 Speed of Sound in Various Bulk Materials

Material	Speed of Sound (m/s)
Gases	
Air (20°C)	344
Helium (20°C)	999
Hydrogen (20°C)	1330
Liquids	
Liquid helium (4 K)	211
Mercury (20°C)	1451
Water (0°C)	1402
Water (20°C)	1482
Water (100°C)	1543
Solids	
Aluminum	6420
Lead	1960
Steel	5941

EXAMPLE 16.3 Wavelength of sonar waves

A ship uses a sonar system (**Fig. 16.8**) to locate underwater objects. Find the speed of sound waves in water using Eq. (16.7), and find the wavelength of a 262 Hz wave.

IDENTIFY and SET UP Our target variables are the speed and wavelength of a sound wave in water. In Eq. (16.7), we use the density of water, $\rho = 1.00 \times 10^3\ \text{kg/m}^3$, and the bulk modulus of water, which we find from the compressibility (see Table 11.2). Given the speed and the frequency $f = 262$ Hz, we find the wavelength from $v = f\lambda$.

EXECUTE In Example 16.2, we used Table 11.2 to find $B = 2.18 \times 10^9$ Pa. Then

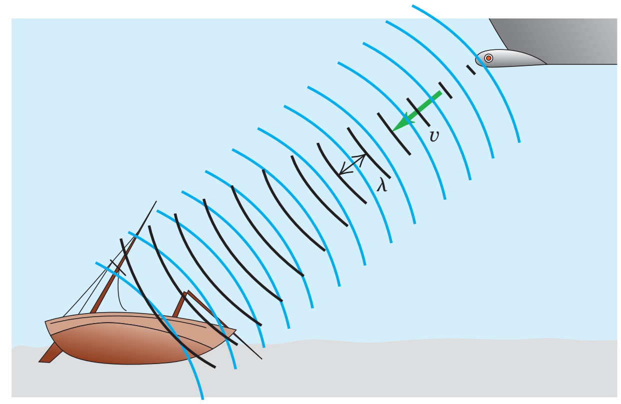
$$v = \sqrt{\frac{B}{\rho}} = \sqrt{\frac{2.18 \times 10^9\ \text{Pa}}{1.00 \times 10^3\ \text{kg/m}^3}} = 1476\ \text{m/s}$$

and

$$\lambda = \frac{v}{f} = \frac{1476\ \text{m/s}}{262\ \text{s}^{-1}} = 5.64\ \text{m}$$

EVALUATE The calculated value of v agrees well with the value in Table 16.1. Water is denser than air (ρ is larger) but is also much more incompressible (B is much larger), and so the speed $v = \sqrt{B/\rho}$ is greater than the 344 m/s speed of sound in air at ordinary temperatures. The relationship $\lambda = v/f$ then says that a sound wave in water must have a longer wavelength than a wave of the same frequency in air. Indeed, we found in Example 15.1 (Section 15.2) that a 262 Hz sound wave in air has a wavelength of only 1.31 m.

Figure 16.8 A sonar system uses underwater sound waves to detect and locate submerged objects.



KEYCONCEPT The speed of sound waves in a fluid depends on the fluid's bulk modulus and density. A sound wave of a given frequency has a longer wavelength in a medium that has a faster sound speed.

Figure 16.9 This three-dimensional image of a fetus in the womb was made using a sequence of ultrasound scans. Each individual scan reveals a two-dimensional "slice" through the fetus; many such slices were then combined digitally. Ultrasound imaging is also used to study heart valve action and to detect tumors.

Dolphins emit high-frequency sound waves (typically 100,000 Hz) and use the echoes for guidance and for hunting. The corresponding wavelength in water is 1.48 cm. With this high-frequency "sonar" system they can sense objects that are roughly as small as the wavelength (but not much smaller). *Ultrasonic imaging* in medicine uses the same principle; sound waves of very high frequency and very short wavelength, called *ultrasound,* are scanned over the human body, and the "echoes" from interior organs are used to create an image. With ultrasound of frequency 5 MHz $= 5 \times 10^6$ Hz, the wavelength in water (the primary constituent of the body) is 0.3 mm, and features as small as this can be discerned in the image (**Fig. 16.9**). Ultrasound is more sensitive than x rays in distinguishing various kinds of tissues and does not have the radiation hazards associated with x rays.

Speed of Sound in a Gas

Most of the sound waves that we encounter propagate in air. To use Eq. (16.7) to find the speed of sound waves in air, we note that the bulk modulus of a gas depends on pressure: The greater the pressure applied to compress a gas, the more it resists further compression and hence the greater the bulk modulus. (That's why specific values of the bulk modulus for gases are not given in Table 11.1.) The expression for the bulk modulus of a gas for use in Eq. (16.7) is

$$B = \gamma p_0 \tag{16.9}$$

where p_0 is the equilibrium pressure of the gas. The quantity γ (the Greek letter gamma) is called the *ratio of heat capacities.* It is a dimensionless number that characterizes the thermal properties of the gas. (We'll learn more about this quantity in Chapter 19.) As an example, the ratio of heat capacities for air is $\gamma = 1.40$. At normal atmospheric pressure $p_0 = 1.013 \times 10^5$ Pa, so $B = (1.40)(1.013 \times 10^5\ \text{Pa}) = 1.42 \times 10^5$ Pa. This value is minuscule compared to the bulk modulus of a typical solid (see Table 11.1), which is approximately 10^{10} to 10^{11} Pa. This shouldn't be surprising: It's simply a statement that air is far easier to compress than steel.

The density ρ of a gas also depends on the pressure, which in turn depends on the temperature. It turns out that the ratio B/ρ for a given type of ideal gas does *not* depend on the pressure at all, only the temperature. From Eq. (16.7), this means that the speed of sound in a gas is fundamentally a function of temperature T:

Speed of sound in an ideal gas

$$v = \sqrt{\frac{\gamma RT}{M}} \tag{16.10}$$

Ratio of heat capacities (γ); Gas constant (R); Absolute temperature (T); Molar mass (M)

This expression incorporates several quantities that we'll study in Chapters 17, 18, and 19. The temperature T is the *absolute* temperature in kelvins (K), equal to the Celsius temperature plus 273.15; thus 20.00°C corresponds to $T = 293.15$ K. The quantity M is the *molar mass,* or mass per mole of the substance of which the gas is composed. The *gas constant* R has the same value for all gases. The current best numerical value of R is

$$R = 8.3144598(48)\ \text{J/mol} \cdot \text{K}$$

which for practical calculations we can write as 8.314 J/mol · K.

For any particular gas, γ, R, and M are constants, and the wave speed is proportional to the square root of the absolute temperature. We'll see in Chapter 18 that Eq. (16.10) is almost identical to the expression for the average speed of molecules in an ideal gas. This shows that sound speeds and molecular speeds are closely related.

EXAMPLE 16.4 Speed of sound in air

Find the speed of sound in air at $T = 20°\text{C}$, and find the range of wavelengths in air to which the human ear (which can hear frequencies in the range of 20–20,000 Hz) is sensitive. The mean molar mass for air (a mixture of mostly nitrogen and oxygen) is $M = 28.8 \times 10^{-3}$ kg/mol and the ratio of heat capacities is $\gamma = 1.40$.

IDENTIFY and SET UP We use Eq. (16.10) to find the sound speed from γ, T, and M, and we use $v = f\lambda$ to find the wavelengths corresponding to the frequency limits. Note that in Eq. (16.10) temperature T *must* be expressed in kelvins, not Celsius degrees.

EXECUTE At $T = 20°\text{C} = 293$ K, we find

$$v = \sqrt{\frac{\gamma RT}{M}} = \sqrt{\frac{(1.40)(8.314\ \text{J/mol}\cdot\text{K})(293\ \text{K})}{28.8 \times 10^{-3}\ \text{kg/mol}}} = 344\ \text{m/s}$$

Using this value of v in $\lambda = v/f$, we find that at 20°C the frequency $f = 20$ Hz corresponds to $\lambda = 17$ m and $f = 20{,}000$ Hz to $\lambda = 1.7$ cm.

EVALUATE Our calculated value of v agrees with the measured sound speed at $T = 20°\text{C}$.

KEYCONCEPT The speed of sound in a gas is determined by the temperature of the gas, its molar mass, and its ratio of heat capacities.

A gas is actually composed of molecules in random motion, separated by distances that are large in comparison with their diameters. The vibrations that constitute a wave in a gas are superposed on the random thermal motion. At atmospheric pressure, a molecule travels an average distance of about 10^{-7} m between collisions, while the displacement amplitude of a faint sound may be only 10^{-9} m. We can think of a gas with a sound wave passing through as being comparable to a swarm of bees; the swarm as a whole oscillates slightly while individual insects move about within the swarm, apparently at random.

TEST YOUR UNDERSTANDING OF SECTION 16.2 Mercury is 13.6 times denser than water. Based on Table 16.1, at 20°C which of these liquids has the greater bulk modulus? (i) Mercury; (ii) water; (iii) both are about the same; (iv) not enough information is given to decide.

ANSWER

(i) From Eq. (16.7), the speed of longitudinal waves (sound) in a fluid is $v = \sqrt{B/\rho}$. We can rewrite this to give an expression for the bulk modulus B in terms of the fluid density ρ and the sound speed v: $B = \rho v^2$. At 20°C the speed of sound in mercury is slightly less than in water (1451 m/s versus 1482 m/s), but the density of mercury is greater than that of water by a large factor (13.6). Hence the bulk modulus of mercury is greater than that of water by a factor of $(13.6)(1451/1482)^2 = 13.0$.

16.3 SOUND INTENSITY

Traveling sound waves, like all other traveling waves, transfer energy from one region of space to another. In Section 15.5 we introduced the *wave intensity* I, equal to the time average rate at which wave energy is transported per unit area across a surface perpendicular to the direction of propagation. Let's see how to express the intensity of a sound wave in a fluid in terms of the displacement amplitude A or pressure amplitude $p_{\max}$.

Let's consider a sound wave propagating in the $+x$-direction so that we can use our expressions from Section 16.1 for the displacement $y(x, t)$ [Eq. (16.1)] and pressure fluctuation $p(x, t)$ [Eq. (16.4)]. In Section 6.4 we saw that power equals the product of force and velocity [see Eq. (6.18)]. So the power per unit area in this sound wave equals the product of $p(x, t)$ (force per unit area) and the *particle* velocity $v_y(x, t)$, which is the velocity at time t of that portion of the wave medium at coordinate x. Using Eqs. (16.1) and (16.4), we find

$$v_y(x, t) = \frac{\partial y(x, t)}{\partial t} = \omega A \sin(kx - \omega t)$$

$$\begin{aligned} p(x, t)v_y(x, t) &= [BkA \sin(kx - \omega t)][\omega A \sin(kx - \omega t)] \\ &= B\omega k A^2 \sin^2(kx - \omega t) \end{aligned}$$

CAUTION **Wave velocity vs. particle velocity** Remember that the velocity of the wave as a whole is *not* the same as the particle velocity. While the wave continues to move in the direction of propagation, individual particles in the wave medium merely slosh back and forth, as shown in Fig. 16.1. Furthermore, the maximum speed of a particle of the medium can be very different from the wave speed.

The intensity is the time average value of the power per unit area $p(x, t)v_y(x, t)$. For any value of x the average value of the function $\sin^2(kx - \omega t)$ over one period $T = 2\pi/\omega$ is $\frac{1}{2}$, so

$$I = \tfrac{1}{2} B\omega k A^2 \tag{16.11}$$

Using the relationships $\omega = vk$ and $v = \sqrt{B/\rho}$, we can rewrite Eq. (16.11) as

Intensity of a sinusoidal sound wave in a fluid

$$I = \tfrac{1}{2}\sqrt{\rho B}\,\omega^2 A^2 \tag{16.12}$$

Angular frequency $= 2\pi f$; Displacement amplitude; Density of fluid; Bulk modulus of fluid

It is usually more useful to express I in terms of the pressure amplitude $p_{\max}$. Using Eqs. (16.5) and (16.12) and the relationship $\omega = vk$, we find

$$I = \frac{\omega p_{\max}{}^2}{2Bk} = \frac{v p_{\max}{}^2}{2B} \tag{16.13}$$

By using the wave speed relationship $v = \sqrt{B/\rho}$, we can also write Eq. (16.13) in the alternative forms

Intensity of a sinusoidal sound wave in a fluid

$$I = \frac{p_{\max}{}^2}{2\rho v} = \frac{p_{\max}{}^2}{2\sqrt{\rho B}} \tag{16.14}$$

Pressure amplitude; Wave speed; Density of fluid; Bulk modulus of fluid

You should verify these expressions. Comparison of Eqs. (16.12) and (16.14) shows that sinusoidal sound waves of the same intensity but different frequency have different displacement amplitudes A but the *same* pressure amplitude $p_{\max}$. This is another reason it is usually more convenient to describe a sound wave in terms of pressure fluctuations, not displacement.

The *total* average power carried across a surface by a sound wave equals the product of the intensity at the surface and the surface area, if the intensity over the surface is uniform. The total average sound power emitted by a person speaking in an ordinary conversational tone is about 10^{-5} W, while a loud shout corresponds to about 3×10^{-2} W. If all the residents of New York City were to talk at the same time, the total sound power would be about 100 W, equivalent to the electric power requirement of a medium-sized light bulb. On the other hand, the power required to fill a large auditorium or stadium with loud sound is considerable (see Example 16.7).

If the sound source emits waves in all directions equally, the intensity decreases with increasing distance r from the source according to the inverse-square law (Section 15.5): The intensity is proportional to $1/r^2$. The intensity can be increased by confining the sound waves to travel in the desired direction only (**Fig. 16.10**), although the $1/r^2$ law still applies.

The inverse-square relationship also does not apply indoors because sound energy can reach a listener by reflection from the walls and ceiling. Indeed, part of the architect's job in designing an auditorium is to tailor these reflections so that the intensity is as nearly uniform as possible over the entire auditorium.

Figure **16.10** By cupping your hands like this, you direct the sound waves emerging from your mouth so that they don't propagate to the sides. Hence you can be heard at greater distances.

PROBLEM-SOLVING STRATEGY 16.1 Sound Intensity

IDENTIFY *the relevant concepts:* The relationships between the intensity and amplitude of a sound wave are straightforward. Other quantities are involved in these relationships, however, so it's particularly important to decide which is your target variable.

SET UP *the problem* using the following steps:

1. Sort the physical quantities into categories. Wave properties include the displacement and pressure amplitudes A and p_{max}. The frequency f can be determined from the angular frequency ω, the wave number k, or the wavelength λ. These quantities are related through the wave speed v, which is determined by properties of the medium (B and ρ for a liquid, and γ, T, and M for a gas).
2. List the given quantities and identify the target variables. Find relationships that take you where you want to go.

EXECUTE *the solution:* Use your selected equations to solve for the target variables. Express the temperature in kelvins (Celsius temperature plus 273.15) to calculate the speed of sound in a gas.

EVALUATE *your answer:* If possible, use an alternative relationship to check your results.

EXAMPLE 16.5 Intensity of a sound wave in air

WITH VARIATION PROBLEMS

Find the intensity of the sound wave in Example 16.1, with $p_{max} = 3.0 \times 10^{-2}$ Pa. Assume the temperature is 20°C so that the density of air is $\rho = 1.20\ \text{kg/m}^3$ and the speed of sound is $v = 344$ m/s.

IDENTIFY and SET UP Our target variable is the intensity I of the sound wave. We are given the pressure amplitude p_{max} of the wave as well as the density ρ and wave speed v for the medium. We can determine I from p_{max}, ρ, and v from Eq. (16.14).

EXECUTE From Eq. (16.14),

$$I = \frac{p_{max}^2}{2\rho v} = \frac{(3.0 \times 10^{-2}\ \text{Pa})^2}{2(1.20\ \text{kg/m}^3)(344\ \text{m/s})}$$
$$= 1.1 \times 10^{-6}\ \text{J/(s} \cdot \text{m}^2) = 1.1 \times 10^{-6}\ \text{W/m}^2$$

EVALUATE This seems like a very low intensity, but it is well within the range of sound intensities encountered on a daily basis. A very loud sound wave at the threshold of pain has a pressure amplitude of about 30 Pa and an intensity of about 1 W/m^2. The pressure amplitude of the faintest sound wave that can be heard is about 3×10^{-5} Pa, and the corresponding intensity is about 10^{-12} W/m^2. (Try these values of p_{max} in Eq. (16.14) to check that the corresponding intensities are as we have stated.)

KEYCONCEPT The intensity (power per unit area) of a sound wave is proportional to the square of the pressure amplitude of the wave. The proportionality constant depends on the density of the medium and the speed of sound in the medium.

EXAMPLE 16.6 Same intensity, different frequencies

WITH VARIATION PROBLEMS

What are the pressure and displacement amplitudes of a 20 Hz sound wave with the same intensity as the 1000 Hz sound wave of Examples 16.1 and 16.5?

IDENTIFY and SET UP In Examples 16.1 and 16.5 we found that for a 1000 Hz sound wave with $p_{max} = 3.0 \times 10^{-2}$ Pa, $A = 1.2 \times 10^{-8}$ m and $I = 1.1 \times 10^{-6}\ \text{W/m}^2$. Our target variables are p_{max} and A for a 20 Hz sound wave of the same intensity I. We can find these using Eqs. (16.14) and (16.12), respectively.

EXECUTE We can rearrange Eqs. (16.14) and (16.12) as $p_{max}{}^2 = 2I\sqrt{\rho B}$ and $\omega^2 A^2 = 2I/\sqrt{\rho B}$, respectively. These tell us that for a given sound intensity I in a given medium (constant ρ and B), the quantities p_{max} and ωA (or, equivalently, fA) are *constants* that don't depend on frequency. From the first result we immediately have $p_{max} = 3.0 \times 10^{-2}$ Pa for $f = 20$ Hz, the same as for $f = 1000$ Hz. If we write the second result as $f_{20}A_{20} = f_{1000}A_{1000}$, we have

$$A_{20} = \left(\frac{f_{1000}}{f_{20}}\right)A_{1000}$$
$$= \left(\frac{1000\ \text{Hz}}{20\ \text{Hz}}\right)(1.2 \times 10^{-8}\ \text{m}) = 6.0 \times 10^{-7}\ \text{m} = 0.60\ \mu\text{m}$$

EVALUATE Our result reinforces the idea that pressure amplitude is a more convenient description of a sound wave and its intensity than displacement amplitude.

KEYCONCEPT If two sound waves in a given medium have the same intensity but different frequencies, the wave with the higher frequency has the greater *displacement* amplitude. The two waves have the same *pressure* amplitude, however.

EXAMPLE 16.7 "Play it loud!"

WITH VARIATION PROBLEMS

For an outdoor concert we want the sound intensity to be $1\ \text{W/m}^2$ at a distance of 20 m from the speaker array. If the sound intensity is uniform in all directions, what is the required average acoustic power output of the array?

IDENTIFY, SET UP, and EXECUTE This example uses the definition of sound intensity as power per unit area. The total power is the target variable; the area in question is a hemisphere centered on the speaker array. We assume that the speakers are on the ground and that none of the acoustic power is directed into the ground, so the acoustic power is uniform over a hemisphere 20 m in radius. The surface area of this hemisphere is $\left(\frac{1}{2}\right)(4\pi)(20\ \text{m})^2$, or about $2500\ \text{m}^2$. The required power is the product of this area and the intensity: $(1\ \text{W/m}^2)(2500\ \text{m}^2) = 2500\ \text{W} = 2.5$ kW.

EVALUATE The electrical power input to the speaker would need to be considerably greater than 2.5 kW, because speaker efficiency is not very high (typically a few percent for ordinary speakers, and up to 25% for horn-type speakers).

KEYCONCEPT To find the acoustic power output of a source of sound, multiply the area over which the emitted sound wave is distributed by the average intensity of the sound over that area.

The Decibel Scale

Because the ear is sensitive over a broad range of intensities, a *logarithmic* measure of intensity called **sound intensity level** is often used:

Sound intensity level; Intensity of sound wave; Reference intensity $= 10^{-12}\ \text{W/m}^2$; Logarithm to base 10

$$\beta = (10\ \text{dB}) \log \frac{I}{I_0} \qquad (16.15)$$

The chosen reference intensity I_0 in Eq. (16.15) is approximately the threshold of human hearing at 1000 Hz. Sound intensity levels are expressed in **decibels,** abbreviated dB. A decibel is $\frac{1}{10}$ of a *bel,* a unit named for Alexander Graham Bell (the inventor of the telephone). The bel is inconveniently large for most purposes, and the decibel is the usual unit of sound intensity level.

If the intensity of a sound wave equals I_0 or $10^{-12}\ \text{W/m}^2$, its sound intensity level is $\beta = 0$ dB. An intensity of $1\ \text{W/m}^2$ corresponds to 120 dB. **Table 16.2** gives the sound intensity levels of some familiar sounds. You can use Eq. (16.15) to check the value of β given for each intensity in the table.

Because the ear is not equally sensitive to all frequencies in the audible range, some sound-level meters weight the various frequencies unequally. One such scheme leads to the so-called dBA scale; this scale deemphasizes the low and very high frequencies, where the ear is less sensitive.

TABLE 16.2 Sound Intensity Levels from Various Sources (Representative Values)

Source or Description of Sound	Sound Intensity Level, β (dB)	Intensity, I (W/m^2)
Military jet aircraft 30 m away	140	10^2
Threshold of pain	120	1
Riveter	95	3.2×10^{-3}
Elevated train	90	10^{-3}
Busy street traffic	70	10^{-5}
Ordinary conversation	65	3.2×10^{-6}
Quiet automobile	50	10^{-7}
Quiet radio in home	40	10^{-8}
Average whisper	20	10^{-10}
Rustle of leaves	10	10^{-11}
Threshold of hearing at 1000 Hz	0	10^{-12}

EXAMPLE 16.8 Temporary—or permanent—hearing loss

WITH VARIATION PROBLEMS

A 10 min exposure to 120 dB sound will temporarily shift your threshold of hearing at 1000 Hz from 0 dB up to 28 dB. Ten years of exposure to 92 dB sound will cause a *permanent* shift to 28 dB. What sound intensities correspond to 28 dB and 92 dB?

IDENTIFY and SET UP We are given two sound intensity levels β; our target variables are the corresponding intensities. We can solve Eq. (16.15) to find the intensity I that corresponds to each value of β.

EXECUTE We solve Eq. (16.15) for I by dividing both sides by 10 dB and using the relationship $10^{\log x} = x$:

$$I = I_0 10^{\beta/(10\text{ dB})}$$

For $\beta = 28$ dB and $\beta = 92$ dB, the exponents are $\beta/(10\text{ dB}) = 2.8$ and 9.2, respectively, so that

$$I_{28\text{ dB}} = (10^{-12}\text{ W/m}^2)10^{2.8} = 6.3 \times 10^{-10}\text{ W/m}^2$$
$$I_{92\text{ dB}} = (10^{-12}\text{ W/m}^2)10^{9.2} = 1.6 \times 10^{-3}\text{ W/m}^2$$

EVALUATE If your answers are a factor of 10 too large, you may have entered 10×10^{-12} in your calculator instead of 1×10^{-12}. Be careful!

KEYCONCEPT The sound intensity level (in decibels, or dB) is a logarithmic measure of the intensity of a sound wave. Adding 10 dB to the sound intensity level corresponds to multiplying the intensity by a factor of 10.

EXAMPLE 16.9 A bird sings in a meadow

WITH VARIATION PROBLEMS

Consider an idealized bird (treated as a point source) that emits constant sound power, with intensity obeying the inverse-square law (**Fig. 16.11**). If you move twice the distance from the bird, by how many decibels does the sound intensity level drop?

IDENTIFY and SET UP The decibel scale is logarithmic, so the *difference* between two sound intensity levels (the target variable) corresponds to the *ratio* of the corresponding intensities, which is determined by the inverse-square law. We label the two points P_1 and P_2 (Fig. 16.11). We use Eq. (16.15), the definition of sound intensity level, at each point. We use Eq. (15.26), the inverse-square law, to relate the intensities at the two points.

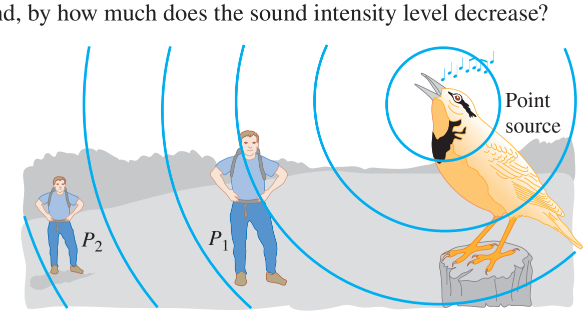
EXECUTE The difference $\beta_2 - \beta_1$ between any two sound intensity levels is related to the corresponding intensities by

$$\beta_2 - \beta_1 = (10\text{ dB})\left(\log\frac{I_2}{I_0} - \log\frac{I_1}{I_0}\right)$$
$$= (10\text{ dB})[(\log I_2 - \log I_0) - (\log I_1 - \log I_0)]$$
$$= (10\text{ dB})\log\frac{I_2}{I_1}$$

For this inverse-square-law source, Eq. (15.26) yields $I_2/I_1 = r_1^2/r_2^2 = \frac{1}{4}$, so

$$\beta_2 - \beta_1 = (10\text{ dB})\log\frac{I_2}{I_1} = (10\text{ dB})\log\tfrac{1}{4} = -6.0\text{ dB}$$

Figure 16.11 When you double your distance from a point source of sound, by how much does the sound intensity level decrease?



EVALUATE Our result is negative, which tells us (correctly) that the sound intensity level is less at P_2 than at P_1. The 6 dB difference doesn't depend on the sound intensity level at P_1; *any* doubling of the distance from an inverse-square-law source reduces the sound intensity level by 6 dB.

Note that the perceived *loudness* of a sound is not directly proportional to its intensity. For example, most people interpret an increase of 8 dB to 10 dB in sound intensity level (corresponding to increasing intensity by a factor of 6 to 10) as a doubling of loudness.

KEYCONCEPT The *difference* between the sound intensity levels of two sounds is proportional to the logarithm of the *ratio* of the intensities of those sounds.

TEST YOUR UNDERSTANDING OF SECTION 16.3 You double the intensity of a sound wave in air while leaving the frequency unchanged. (The pressure, density, and temperature of the air remain unchanged as well.) What effect does this have on the displacement amplitude, pressure amplitude, bulk modulus, sound speed, and sound intensity level?

ANSWER

❙ **A and p_{max} increase by a factor of $\sqrt{2}$, B and v are unchanged, β increases by 3.0 dB** Equations (16.9) and (16.10) show that the bulk modulus B and sound speed v remain the same because the physical properties of the air are unchanged. From Eqs. (16.12) and (16.14), the intensity is proportional to the square of the displacement amplitude or the square of the pressure amplitude. Hence doubling the intensity means that A and p_{max} both increase by a factor of $\sqrt{2}$. Example 16.9 shows that *multiplying* the intensity by a factor of 2 $(I_2/I_1 = 2)$ corresponds to *adding* to the sound intensity level by $(10\ \text{dB})\log(I_2/I_1) = (10\ \text{dB})\log 2 = 3.0\ \text{dB}$.

16.4 STANDING SOUND WAVES AND NORMAL MODES

When longitudinal (sound) waves propagate in a fluid in a pipe, the waves are reflected from the ends in the same way that transverse waves on a string are reflected at its ends. The superposition of the waves traveling in opposite directions again forms a standing wave. Just as for transverse standing waves on a string (see Section 15.7), standing sound waves in a pipe can be used to create sound waves in the surrounding air. This is the principle of the human voice as well as many musical instruments, including woodwinds, brasses, and pipe organs.

Transverse waves on a string, including standing waves, are usually described only in terms of the displacement of the string. But, as we have seen, sound waves in a fluid may be described either in terms of the displacement of the fluid or in terms of the pressure variation in the fluid. To avoid confusion, we'll use the terms **displacement node** and **displacement antinode** to refer to points where particles of the fluid have zero displacement and maximum displacement, respectively.

We can demonstrate standing sound waves in a column of gas using an apparatus called a Kundt's tube (**Fig. 16.12**). A horizontal glass tube a meter or so long is closed at one end and has a flexible diaphragm at the other end that can transmit vibrations. A nearby loudspeaker is driven by an audio oscillator and amplifier; this produces sound waves that force the diaphragm to vibrate sinusoidally with a frequency that we can vary. The sound waves within the tube are reflected at the other, closed end of the tube. We spread a small amount of light powder uniformly along the bottom of the tube. As we vary the frequency of the sound, we pass through frequencies at which the amplitude of the standing waves becomes large enough for the powder to be swept along the tube at those points where the gas is in motion. The powder therefore collects at the displacement nodes (where the gas is not moving). Adjacent nodes are separated by a distance equal to $\lambda/2$.

Figure 16.12 Demonstrating standing sound waves using a Kundt's tube. The blue shading represents the density of the gas at an instant when the gas pressure at the displacement nodes is a maximum or a minimum.

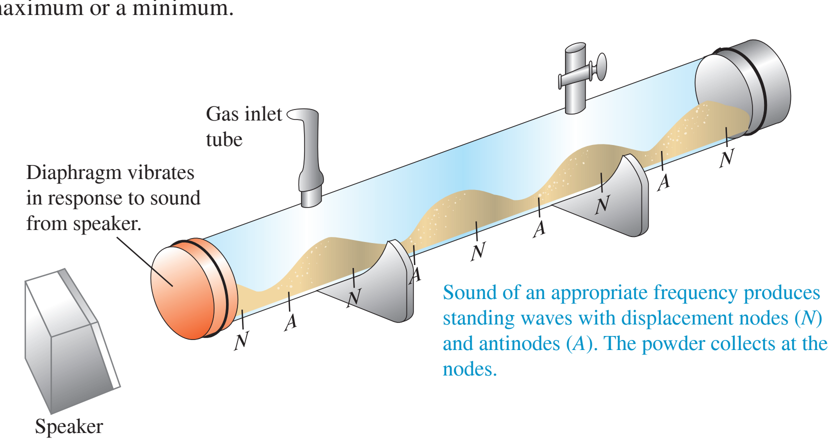


Figure 16.13 shows the motions of nine different particles within a gas-filled tube in which there is a standing sound wave. A particle at a displacement node (N) does not move, while a particle at a displacement antinode (A) oscillates with maximum amplitude. Note that particles on opposite sides of a displacement node vibrate in opposite phase. When these particles approach each other, the gas between them is compressed and the pressure rises; when they recede from each other, there is an expansion and the pressure drops. Hence at a displacement *node* the gas undergoes the maximum amount of compression and expansion, and the variations in pressure and density above and below the average have their maximum value. By contrast, particles on opposite sides of a displacement *antinode* vibrate *in phase;* the distance between the particles is nearly constant, and there is *no* variation in pressure or density at a displacement antinode.

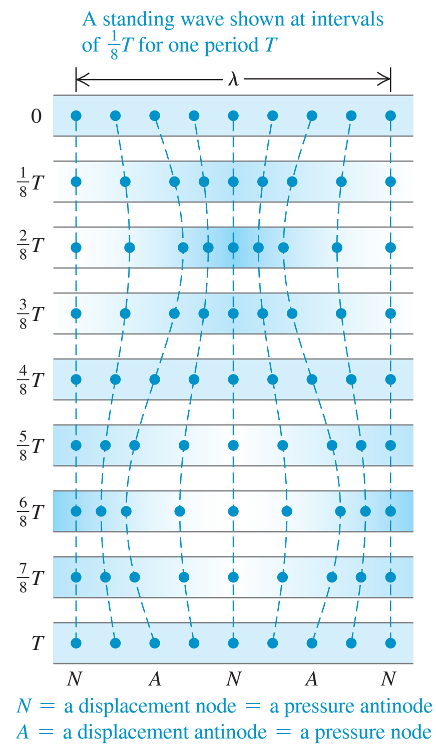
We use the term **pressure node** to describe a point in a standing sound wave at which the pressure and density do not vary and the term **pressure antinode** to describe a point at which the variations in pressure and density are greatest. Using these terms, we can summarize our observations as follows:

A pressure node is always a displacement antinode, and a pressure antinode is always a displacement node.

Figure 16.12 depicts a standing sound wave at an instant at which the pressure variations are greatest; the blue shading shows that the density and pressure of the gas have their maximum and minimum values at the displacement nodes.

When reflection takes place at a *closed* end of a pipe (an end with a rigid barrier or plug), the displacement of the particles at this end must always be zero, analogous to a fixed end of a string. Thus a closed end of a pipe is a displacement node and a pressure antinode; the particles do not move, but the pressure variations are maximum. An *open* end of a pipe is a pressure node because it is open to the atmosphere, where the pressure is constant. Because of this, an open end is always a displacement *antinode,* in analogy to a free end of a string; the particles oscillate with maximum amplitude, but the pressure does not vary. (The pressure node actually occurs somewhat beyond an open end of a pipe. But if the diameter of the pipe is small in comparison to the wavelength, which is true for most musical instruments, this effect can safely be ignored.) Thus longitudinal sound waves are reflected at the closed and open ends of a pipe in the same way that transverse waves in a string are reflected at fixed and free ends, respectively.

Figure 16.13 In a standing sound wave, a displacement node N is a pressure antinode (a point where the pressure fluctuates the most) and a displacement antinode A is a pressure node (a point where the pressure does not fluctuate at all).



CONCEPTUAL EXAMPLE 16.10 The sound of silence

A directional loudspeaker directs a sound wave of wavelength λ at a wall (**Fig. 16.14**). At what distances from the wall could you stand and hear no sound at all?

SOLUTION Your ear detects pressure variations in the air; you'll therefore hear no sound if your ear is at a *pressure node,* which is a displacement antinode. The wall is at a displacement node; the distance from any node to an adjacent antinode is $\lambda/4$, and the distance from one antinode to the next is $\lambda/2$ (Fig. 16.14). Hence the displacement antinodes (pressure nodes), at which no sound will be heard, are at distances $d = \lambda/4$, $d = \lambda/4 + \lambda/2 = 3\lambda/4$, $d = 3\lambda/4 + \lambda/2 = 5\lambda/4, \ldots$ from the wall. If the loudspeaker is not highly directional, this effect is hard to notice because of reflections of sound waves from the floor, ceiling, and other walls.

KEYCONCEPT In a standing sound wave, a pressure node is a displacement antinode, and vice versa. The sound is loudest at a pressure antinode; there is no sound at a pressure node.

Figure 16.14 When a sound wave is directed at a wall, it interferes with the reflected wave to create a standing wave. The N's and A's are *displacement* nodes and antinodes.

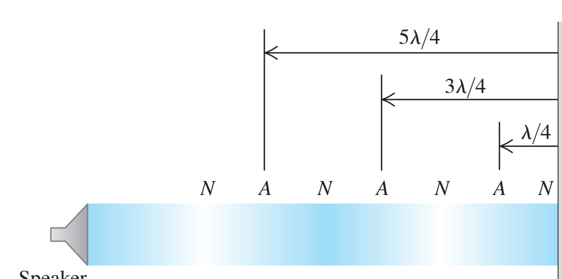


Figure **16.15** Organ pipes of different sizes produce tones with different frequencies.

Organ Pipes and Wind Instruments

The most important application of standing sound waves is the production of musical tones. Organ pipes are one of the simplest examples (**Fig. 16.15**). Air is supplied by a blower to the bottom end of the pipe (**Fig. 16.16**). A stream of air emerges from the narrow opening at the edge of the horizontal surface and is directed against the top edge of the opening, which is called the *mouth* of the pipe. The column of air in the pipe is set into vibration, and there is a series of possible normal modes, just as with the stretched string. The mouth acts as an open end; it is a pressure node and a displacement antinode. The other end of the pipe (at the top in Fig. 16.16) may be either open or closed.

In **Fig. 16.17**, both ends of the pipe are open, so both ends are pressure nodes and displacement antinodes. An organ pipe that is open at both ends is called an *open pipe*. The fundamental frequency f_1 corresponds to a standing-wave pattern with a displacement antinode at each end and a displacement node in the middle (Fig. 16.17a). The distance between adjacent antinodes is always equal to one half-wavelength, and in this case that is equal to the length L of the pipe; $\lambda/2 = L$. The corresponding frequency, obtained from the relationship $f = v/\lambda$, is

$$f_1 = \frac{v}{2L} \qquad \text{(open pipe)} \tag{16.16}$$

Figures 16.17b and 16.17c show the second and third harmonics; their vibration patterns have two and three displacement nodes, respectively. For these, a half-wavelength is equal to $L/2$ and $L/3$, respectively, and the frequencies are twice and three times the fundamental, respectively: $f_2 = 2f_1$ and $f_3 = 3f_1$. For *every* normal mode of an open pipe the length L must be an integer number of half-wavelengths, and the possible wavelengths λ_n are given by

$$L = n\frac{\lambda_n}{2} \quad \text{or} \quad \lambda_n = \frac{2L}{n} \qquad (n = 1, 2, 3, \ldots) \qquad \text{(open pipe)} \tag{16.17}$$

The corresponding frequencies f_n are given by $f_n = v/\lambda_n$, so all the normal-mode frequencies for a pipe that is open at both ends are given by

Standing waves, open pipe:

Frequency of nth harmonic ($n = 1, 2, 3, \ldots$)

$$f_n = \frac{nv}{2L} \tag{16.18}$$

Speed of sound in pipe
Length of pipe

The value $n = 1$ corresponds to the fundamental frequency, $n = 2$ to the second harmonic (or first overtone), and so on. Alternatively, we can say

$$f_n = nf_1 \qquad (n = 1, 2, 3, \ldots) \qquad \text{(open pipe)} \tag{16.19}$$

with f_1 given by Eq. (16.16).

Figure 16.18 shows a *stopped pipe:* It is open at the left end but closed at the right end. The left (open) end is a displacement antinode (pressure node), but the right (closed) end is a displacement node (pressure antinode). Figure 16.18a shows the lowest-frequency

Figure **16.16** Cross sections of an organ pipe at two instants one half-period apart. The N's and A's are *displacement* nodes and antinodes; as the blue shading shows, these are points of maximum pressure variation and zero pressure variation, respectively.

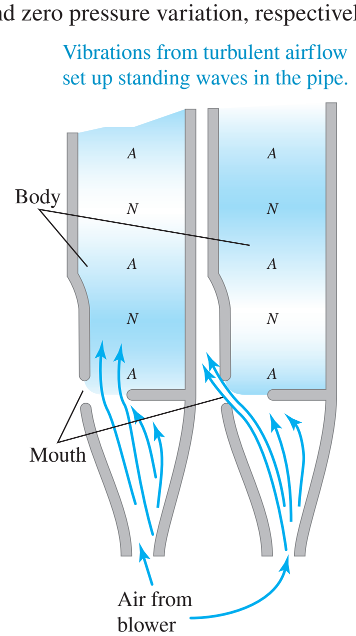


Figure **16.17** A cross section of an open pipe showing the first three normal modes. The shading indicates the pressure variations. The red curves are graphs of the displacement along the pipe axis at two instants separated in time by one half-period. The N's and A's are the *displacement* nodes and antinodes; interchange these to show the *pressure* nodes and antinodes.

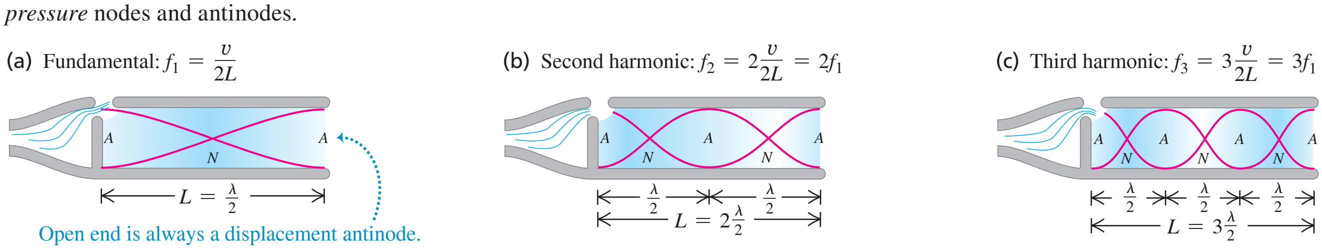
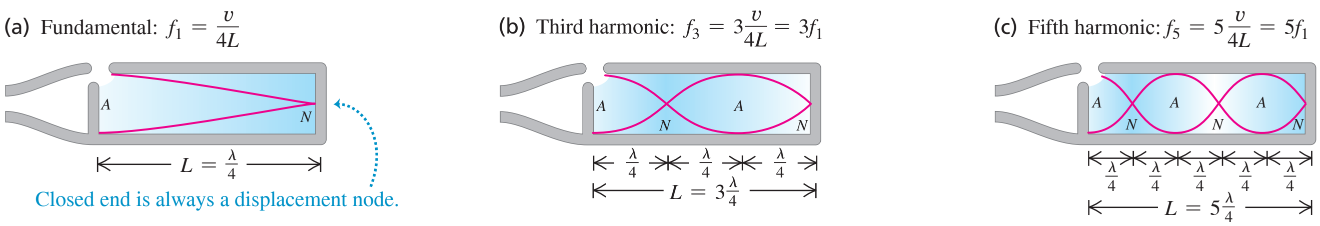


Figure 16.18 A cross section of a stopped pipe showing the first three normal modes as well as the *displacement* nodes and antinodes. Only odd harmonics are possible.



mode; the length of the pipe is the distance between a node and the adjacent antinode, or a quarter-wavelength $(L = \lambda_1/4)$. The fundamental frequency is $f_1 = v/\lambda_1$, or

$$f_1 = \frac{v}{4L} \quad \text{(stopped pipe)} \tag{16.20}$$

This is one-half the fundamental frequency for an *open* pipe of the same length. In musical language, the *pitch* of a closed pipe is one octave lower (a factor of 2 in frequency) than that of an open pipe of the same length. Figure 16.18b shows the next mode, for which the length of the pipe is *three-quarters* of a wavelength, corresponding to a frequency $3f_1$. For Fig. 16.18c, $L = 5\lambda/4$ and the frequency is $5f_1$. The possible wavelengths are given by

$$L = n\frac{\lambda_n}{4} \quad \text{or} \quad \lambda_n = \frac{4L}{n} \quad (n = 1, 3, 5, \ldots) \quad \text{(stopped pipe)} \tag{16.21}$$

The normal-mode frequencies are given by $f_n = v/\lambda_n$, or

Standing waves, stopped pipe:

Frequency of nth harmonic ($n = 1, 3, 5, \ldots$)

$$f_n = \frac{nv}{4L} \tag{16.22}$$

v: Speed of sound in pipe

L: Length of pipe

or

$$f_n = nf_1 \quad (n = 1, 3, 5, \ldots) \quad \text{(stopped pipe)} \tag{16.23}$$

with f_1 given by Eq. (16.20). We see that the second, fourth, and all *even* harmonics are missing. In a stopped pipe, the fundamental frequency is $f_1 = v/4L$, and only the odd harmonics in the series $(3f_1, 5f_1, \ldots)$ are possible.

A final possibility is a pipe that is closed at *both* ends, with displacement nodes and pressure antinodes at both ends. This wouldn't be of much use as a musical instrument because the vibrations couldn't get out of the pipe.

EXAMPLE 16.11 A tale of two pipes

WITH VARIATION PROBLEMS

On a day when the speed of sound is 344 m/s, the fundamental frequency of a particular stopped organ pipe is 220 Hz. (a) How long is this pipe? (b) The second *overtone* of this pipe has the same wavelength as the third *harmonic* of an *open* pipe. How long is the open pipe?

IDENTIFY and SET UP This problem uses the relationship between the length and normal-mode frequencies of open pipes (Fig. 16.17) and stopped pipes (Fig. 16.18). In part (a), we determine the length of the stopped pipe from Eq. (16.22). In part (b), we must determine the length of an open pipe, for which Eq. (16.18) gives the frequencies.

EXECUTE (a) For a stopped pipe $f_1 = v/4L$, so

$$L_{\text{stopped}} = \frac{v}{4f_1} = \frac{344\text{ m/s}}{4(220\text{ s}^{-1})} = 0.391\text{ m}$$

(b) The frequency of the second overtone of a stopped pipe (the *third* possible frequency) is $f_5 = 5f_1 = 5(220\text{ Hz}) = 1100\text{ Hz}$. If the wavelengths for the two pipes are the same, the frequencies are also the same. Hence the frequency of the third harmonic of the open pipe, which is at $3f_1 = 3(v/2L)$, equals 1100 Hz. Then

$$1100\text{ Hz} = 3\left(\frac{344\text{ m/s}}{2L_{\text{open}}}\right) \quad \text{and} \quad L_{\text{open}} = 0.469\text{ m}$$

EVALUATE The 0.391 m stopped pipe has a fundamental frequency of 220 Hz; the *longer* (0.469 m) open pipe has a *higher* fundamental frequency, $(1100\text{ Hz})/3 = 367\text{ Hz}$. This is not a contradiction, as you can see if you compare Figs. 16.17a and 16.18a.

KEYCONCEPT For a pipe open at both ends (an "open pipe"), the normal-mode frequencies of a standing sound wave include both even and odd multiples of the pipe's fundamental frequency. For a pipe open at one end and closed at the other (a "stopped pipe"), the only normal-mode frequencies are the odd multiples of the pipe's fundamental frequency. The fundamental frequency of a stopped pipe is half that of an open pipe of the same length.

In an organ pipe in actual use, several modes are always present at once; the motion of the air is a superposition of these modes. This situation is analogous to a string that is struck or plucked, as in Fig. 15.28. Just as for a vibrating string, a complex standing wave in the pipe produces a traveling sound wave in the surrounding air with a harmonic content similar to that of the standing wave. A very narrow pipe produces a sound wave rich in higher harmonics; a fatter pipe produces mostly the fundamental mode, heard as a softer, more flutelike tone. The harmonic content also depends on the shape of the pipe's mouth.

We have talked about organ pipes, but this discussion is also applicable to other wind instruments. The flute and the recorder are directly analogous. The most significant difference is that those instruments have holes along the pipe. Opening and closing the holes with the fingers changes the effective length L of the air column and thus changes the pitch. Any individual organ pipe, by comparison, can play only a single note. The flute and recorder behave as *open* pipes, while the clarinet acts as a *stopped* pipe (closed at the reed end, open at the bell).

Equations (16.18) and (16.22) show that the frequencies of any wind instrument are proportional to the speed of sound v in the air column inside the instrument. As Eq. (16.10) shows, v depends on temperature; it increases when temperature increases. Thus the pitch of all wind instruments rises with increasing temperature. An organ that has some of its pipes at one temperature and others at a different temperature is bound to sound out of tune.

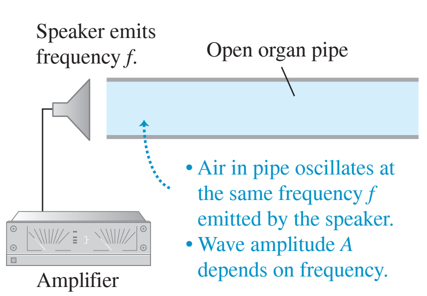
TEST YOUR UNDERSTANDING OF SECTION 16.4 If you connect a hose to one end of a metal pipe and blow compressed air into it, the pipe produces a musical tone. If instead you blow compressed helium into the pipe at the same pressure and temperature, will the pipe produce (i) the same tone, (ii) a higher-pitch tone, or (iii) a lower-pitch tone?

ANSWER **(ii)** Helium is less dense and has a lower molar mass than air, so sound travels faster in helium than in air. The normal-mode frequencies for a pipe are proportional to the sound speed v, so the frequency and hence the pitch increase when the air in the pipe is replaced with helium.

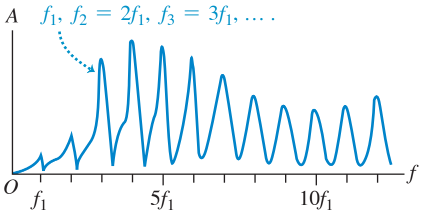
16.5 RESONANCE AND SOUND

Figure 16.19 (a) The air in an open pipe is forced to oscillate at the same frequency as the sinusoidal sound waves coming from the loudspeaker. (b) The resonance curve of the open pipe graphs the amplitude of the standing sound wave in the pipe as a function of the driving frequency.

(a)



(b) Resonance curve: graph of amplitude A versus driving frequency f. Peaks occur at normal-mode frequencies of the pipe: $f_1, f_2 = 2f_1, f_3 = 3f_1, \ldots$.



Many mechanical systems have normal modes of oscillation. As we have seen, these include columns of air (as in an organ pipe) and stretched strings (as in a guitar; see Section 15.8). In each mode, every particle of the system oscillates with simple harmonic motion at the same frequency as the mode. Air columns and stretched strings have an infinite series of normal modes, but the basic concept is closely related to the simple harmonic oscillator, discussed in Chapter 14, which has only a single normal mode (that is, only one frequency at which it oscillates after being disturbed).

Suppose we apply a periodically varying force to a system that can oscillate. The system is then forced to oscillate with a frequency equal to the frequency of the applied force (called the *driving frequency*). This motion is called a *forced oscillation*. We talked about forced oscillations of the harmonic oscillator in Section 14.8, including the phenomenon of mechanical **resonance.** A simple example of resonance is pushing Cousin Throckmorton on a swing. The swing is a pendulum; it has only a single normal mode, with a frequency determined by its length. If we push the swing periodically with this frequency, we can build up the amplitude of the motion. But if we push with a very different frequency, the swing hardly moves at all.

Resonance also occurs when a periodically varying force is applied to a system with many normal modes. In **Fig. 16.19a** an open organ pipe is placed next to a loudspeaker that emits pure sinusoidal sound waves of frequency f, which can be varied by adjusting the amplifier. The air in the pipe is forced to vibrate with the same frequency f as the *driving force* provided by the loudspeaker. In general the amplitude of this motion is relatively small, and the air inside the pipe will not move in any of the normal-mode patterns shown in Fig. 16.17. But if the frequency f of the force is close to one of the normal-mode frequencies, the air in the pipe moves in the normal-mode pattern for that frequency, and the amplitude can become quite large. Figure 16.19b shows the amplitude of oscillation of the air in the pipe as a function of the driving frequency f. This **resonance curve** of the pipe has peaks where f equals the normal-mode frequencies of the pipe. The detailed shape of the resonance curve depends on the geometry of the pipe.

If the frequency of the force is precisely *equal* to a normal-mode frequency, the system is in resonance, and the amplitude of the forced oscillation is maximum. If there were no friction or other energy-dissipating mechanism, a driving force at a normal-mode frequency would continue to add energy to the system, the amplitude would increase indefinitely, and the peaks in the resonance curve of Fig. 16.19b would be infinitely high. But in any real system there is always some dissipation of energy, or damping, as we discussed in Section 14.8; the amplitude of oscillation in resonance may be large, but it cannot be infinite.

The "sound of the ocean" you hear when you put your ear next to a large seashell is due to resonance. The noise of the outside air moving past the seashell is a mixture of sound waves of almost all audible frequencies, which forces the air inside the seashell to oscillate. The seashell behaves like an organ pipe, with a set of normal-mode frequencies; hence the inside air oscillates most strongly at those frequencies, producing the seashell's characteristic sound. To hear a similar phenomenon, uncap a full bottle of your favorite beverage and blow across the open top. The noise is provided by your breath blowing across the top, and the "organ pipe" is the column of air inside the bottle above the surface of the liquid. If you take a drink and repeat the experiment, you'll hear a lower tone because the "pipe" is longer and the normal-mode frequencies are lower.

Resonance also occurs when a stretched string is forced to oscillate (see Section 15.8). Suppose that one end of a stretched string is held fixed while the other is given a transverse sinusoidal motion with small amplitude, setting up standing waves. If the frequency of the driving mechanism is *not* equal to one of the normal-mode frequencies of the string, the amplitude at the antinodes is fairly small. However, if the frequency is equal to any one of the normal-mode frequencies, the string is in resonance, and the amplitude at the antinodes is very much larger than that at the driven end. The driven end is not precisely a node, but it lies much closer to a node than to an antinode when the string is in resonance. The photographs of standing waves in Fig. 15.23 were made this way, with the left end of the string fixed and the right end oscillating vertically with small amplitude.

It is easy to demonstrate resonance with a piano. Push down the damper pedal (the right-hand pedal) so that the dampers are lifted and the strings are free to vibrate, and then sing a steady tone into the piano. When you stop singing, the piano seems to continue to sing the same note. The sound waves from your voice excite vibrations in the strings that have natural frequencies close to the frequencies (fundamental and harmonics) present in the note you sang.

A more spectacular example is a singer breaking a wine glass with her amplified voice. A good-quality wine glass has normal-mode frequencies that you can hear by tapping it. If the singer emits a loud note with a frequency corresponding exactly to one of these normal-mode frequencies, large-amplitude oscillations can build up and break the glass (**Fig. 16.20**).

BIO APPLICATION Resonance and the Sensitivity of the Ear The auditory canal of the human ear (see Fig. 16.4) is an air-filled pipe open at one end and closed at the other (eardrum) end. The canal is about 2.5 cm = 0.025 m long, so it has a resonance at its fundamental frequency $f_1 = v/4L = (344\text{ m/s})/[4(0.025\text{ m})] = 3440\text{ Hz}$. The resonance means that a sound at this frequency produces a strong oscillation of the eardrum. That's why your ear is most sensitive to sounds near 3440 Hz.

Figure **16.20** The frequency of the sound from this trumpet exactly matches one of the normal-mode frequencies of the goblet. The resonant vibrations of the goblet have such large amplitude that the goblet tears itself apart.

EXAMPLE 16.12 An organ–guitar duet

WITH VARIATION PROBLEMS

A stopped organ pipe is sounded near a guitar, causing one of the strings to vibrate with large amplitude. We vary the string tension until we find the maximum amplitude. The string is 80% as long as the pipe. If both pipe and string vibrate at their fundamental frequency, calculate the ratio of the wave speed on the string to the speed of sound in air.

IDENTIFY and SET UP The large response of the string is an example of resonance. It occurs because the organ pipe and the guitar string have the same fundamental frequency. If we let the subscripts a and s stand for the air in the pipe and the string, respectively, the condition for resonance is $f_{1\mathrm{a}} = f_{1\mathrm{s}}$. Equation (16.20) gives the fundamental frequency for a stopped pipe, and Eq. (15.32) gives the fundamental frequency for a guitar string held at both ends. These expressions involve the wave speed in air (v_a) and on the string (v_s) and the lengths of the pipe and string. We are given that $L_\mathrm{s} = 0.80L_\mathrm{a}$; our target variable is the ratio $v_\mathrm{s}/v_\mathrm{a}$.

Continued

EXECUTE From Eqs. (16.20) and (15.32), $f_{1a} = v_a/4L_a$ and $f_{1s} = v_s/2L_s$. These frequencies are equal, so

$$\frac{v_a}{4L_a} = \frac{v_s}{2L_s}$$

Substituting $L_s = 0.80L_a$ and rearranging, we get $v_s/v_a = 0.40$.

EVALUATE As an example, if the speed of sound in air is 344 m/s, the wave speed on the string is $(0.40)(344\ \text{m/s}) = 138\ \text{m/s}$. Note that while the standing waves in the pipe and on the string have the same frequency, they have different *wavelengths* $\lambda = v/f$ because the two media have different wave speeds v. Which standing wave has the greater wavelength?

KEYCONCEPT If you force or drive a mechanical system (such as a guitar string or the air in a pipe) to vibrate at a frequency f, the system will oscillate with maximum amplitude (or *resonate*) if f equals one of the normal-mode frequencies of the system.

TEST YOUR UNDERSTANDING OF SECTION 16.5 A stopped organ pipe of length L has a fundamental frequency of 220 Hz. For which of the following organ pipes will there be a resonance if a tuning fork of frequency 660 Hz is sounded next to the pipe? (There may be more than one correct answer.) (i) A stopped organ pipe of length L; (ii) a stopped organ pipe of length $2L$; (iii) an open organ pipe of length L; (iii) an open organ pipe of length $2L$.

ANSWER

(i) and (iv) There will be a resonance if 660 Hz is one of the pipe's normal-mode frequencies. A stopped organ pipe has normal-mode frequencies that are odd multiples of its fundamental frequency [see Eq. (16.22) and Fig. 16.18]. Hence pipe (i), which has fundamental frequency 220 Hz, also has a normal-mode frequency of $3(220\ \text{Hz}) = 660\ \text{Hz}$. Pipe (ii) has twice the length of pipe (i); from Eq. (16.20), the fundamental frequency of a stopped pipe is inversely proportional to the length, so pipe (ii) has a fundamental frequency of $\left(\frac{1}{2}\right)(220\ \text{Hz}) = 110\ \text{Hz}$. Its other normal-mode frequencies are 330 Hz, 550 Hz, 770 Hz, . . . , so a 660 Hz tuning fork will not cause resonance. Pipe (iii) is an open pipe of the same length as pipe (i), so its fundamental frequency is twice as great as for pipe (i) [compare Eqs. (16.16) and (16.20)], or $2(220\ \text{Hz}) = 440\ \text{Hz}$. Its other normal-mode frequencies are integer multiples of the fundamental frequency [see Eq. (16.19)], or 880 Hz, 1320 Hz, . . . , none of which match the 660 Hz frequency of the tuning fork. Pipe (iv) is also an open pipe but with twice the length of pipe (iii) [see Eq. (16.18)], so its normal-mode frequencies are one-half those of pipe (iii): 220 Hz, 440 Hz, 660 Hz, . . . , so the third harmonic will resonate with the tuning fork.

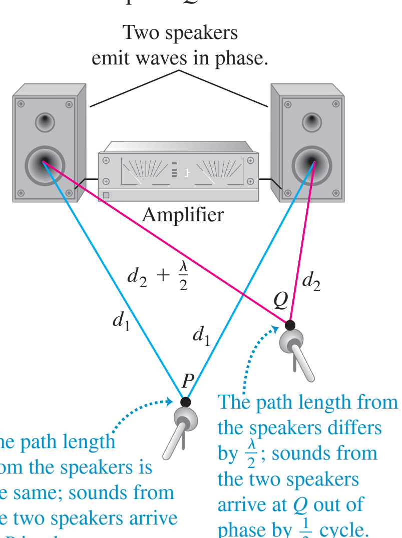
16.6 INTERFERENCE OF WAVES

Wave phenomena that occur when two or more waves overlap in the same region of space are grouped under the heading *interference*. As we have seen, standing waves are a simple example of an interference effect: Two waves traveling in opposite directions in a medium can combine to produce a standing-wave pattern with nodes and antinodes that do not move.

Figure 16.21 shows an example of another type of interference that involves waves that spread out in space. Two speakers, driven in phase by the same amplifier, emit identical sinusoidal sound waves with the same constant frequency. We place a microphone at point P in the figure, equidistant from the speakers. Wave crests emitted from the two speakers at the same time travel equal distances and arrive at point P at the same time; hence the waves arrive in phase, and there is constructive interference. The total wave amplitude that we measure at P is twice the amplitude from each individual wave.

Now let's move the microphone to point Q, where the distances from the two speakers to the microphone differ by a half-wavelength. Then the two waves arrive a half-cycle out of step, or *out of phase;* a positive crest from one speaker arrives at the same time as a negative crest from the other. Destructive interference takes place, and the amplitude measured by the microphone is much *smaller* than when only one speaker is present. If the amplitudes from the two speakers are equal, the two waves cancel each other out completely at point Q, and the total amplitude there is zero.

Figure **16.21** Two speakers driven by the same amplifier. Constructive interference occurs at point P, and destructive interference occurs at point Q.



CAUTION **Interference and traveling waves** The total wave in Fig. 16.21 is a *traveling* wave, not a standing wave. In a standing wave there is no net flow of energy in any direction; by contrast, in Fig. 16.21 there *is* an overall flow of energy from the speakers into the surrounding air, characteristic of a traveling wave. The interference between the waves from the two speakers simply causes the energy flow to be *channeled* into certain directions (for example, toward P) and away from other directions (for example, away from Q). You can see another difference between Fig. 16.21 and a standing wave by considering a point, such as Q, where destructive interference occurs. Such a point is *both* a displacement node *and* a pressure node because there is no wave at all at this point. In a standing wave, a pressure node is a displacement antinode, and vice versa.

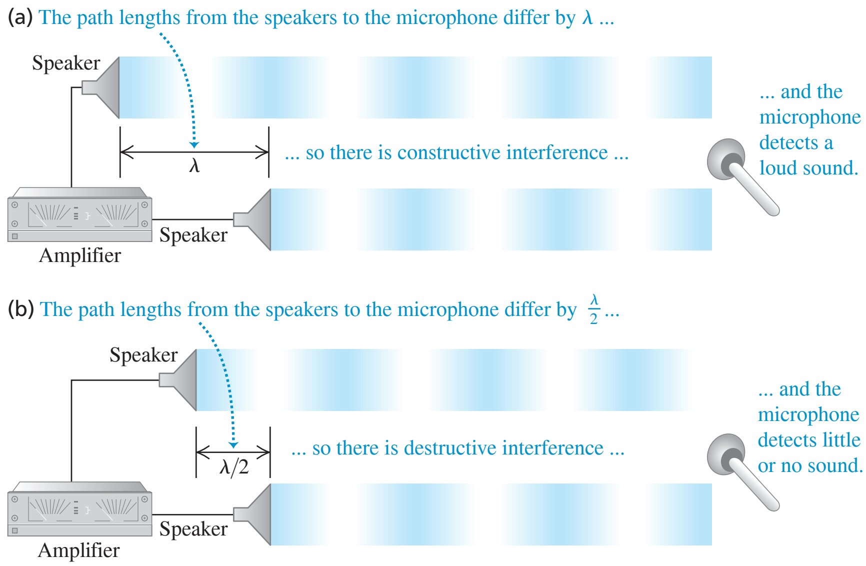


Figure **16.22** Two speakers driven by the same amplifier, emitting waves in phase. Only the waves directed toward the microphone are shown, and they are separated for clarity. (a) Constructive interference occurs when the path difference is $0, \lambda, 2\lambda, 3\lambda, \ldots$. (b) Destructive interference occurs when the path difference is $\lambda/2, 3\lambda/2, 5\lambda/2, \ldots$.

Constructive interference occurs wherever the distances traveled by the two waves differ by a whole number of wavelengths, $0, \lambda, 2\lambda, 3\lambda, \ldots$; then the waves arrive at the microphone in phase (**Fig. 16.22a**). If the distances from the two speakers to the microphone differ by any half-integer number of wavelengths, $\lambda/2, 3\lambda/2, 5\lambda/2, \ldots$, the waves arrive at the microphone out of phase and there will be destructive interference (Fig. 16.22b). In this case, little or no sound energy flows toward the microphone. The energy instead flows in other directions, to where constructive interference occurs.

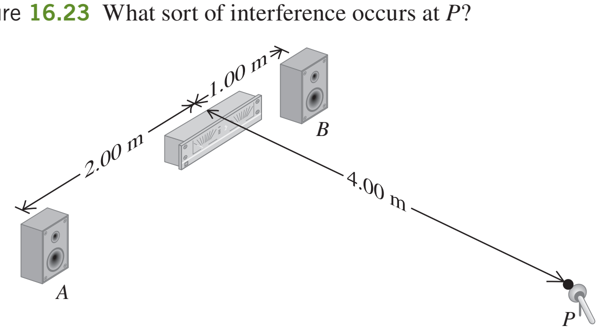
EXAMPLE 16.13 Loudspeaker interference

Two small loudspeakers, A and B (**Fig. 16.23**), are driven by the same amplifier and emit pure sinusoidal waves in phase. (a) For what frequencies does constructive interference occur at point P? (b) For what frequencies does destructive interference occur? The speed of sound is 350 m/s.

IDENTIFY and SET UP The nature of the interference at P depends on the difference d in path lengths from point A to P and from point B to P. We calculate the path lengths using the Pythagorean theorem. Constructive interference occurs when d equals a whole number of wavelengths, while destructive interference occurs when d is a half-integer number of wavelengths. To find the corresponding frequencies, we use $v = f\lambda$.

EXECUTE The A-to-P distance is $[(2.00\text{ m})^2 + (4.00\text{ m})^2]^{1/2} = 4.47\text{ m}$, and the B-to-P distance is $[(1.00\text{ m})^2 + (4.00\text{ m})^2]^{1/2} = 4.12\text{ m}$. The path difference is $d = 4.47\text{ m} - 4.12\text{ m} = 0.35\text{ m}$.

Figure **16.23** What sort of interference occurs at P?



(a) Constructive interference occurs when $d = 0, \lambda, 2\lambda, \ldots$ or $d = 0, v/f, 2v/f, \ldots = nv/f$. So the possible frequencies are

$$f_n = \frac{nv}{d} = n\frac{350\text{ m/s}}{0.35\text{ m}} \qquad (n = 1, 2, 3, \ldots)$$

$$= 1000\text{ Hz}, 2000\text{ Hz}, 3000\text{ Hz}, \ldots$$

(b) Destructive interference occurs when $d = \lambda/2, 3\lambda/2, 5\lambda/2, \ldots$ or $d = v/2f, 3v/2f, 5v/2f, \ldots$. The possible frequencies are

$$f_n = \frac{nv}{2d} = n\frac{350\text{ m/s}}{2(0.35\text{ m})} \qquad (n = 1, 3, 5, \ldots)$$

$$= 500\text{ Hz}, 1500\text{ Hz}, 2500\text{ Hz}, \ldots$$

EVALUATE As we increase the frequency, the sound at point P alternates between large and small (near zero) amplitudes, with maxima and minima at the frequencies given above. This effect may not be strong in an ordinary room because of reflections from the walls, floor, and ceiling.

KEYCONCEPT Two sound waves of the same frequency interfere *constructively* at a certain point if the waves arrive there in phase. If the waves arrive at that point out of phase, they interfere *destructively*.

Figure **16.24** This aviation headset uses destructive interference to minimize the amount of noise from wind and propellers that reaches the wearer's ears.

Interference is the principle behind active noise-reduction headsets, which are used in loud environments such as airplane cockpits (**Fig. 16.24**). A microphone on the headset detects outside noise, and the headset circuitry replays the noise inside the headset shifted in phase by one half-cycle. This phase-shifted sound interferes destructively with the sounds that enter the headset from outside, so the headset wearer experiences very little unwelcome noise.

TEST YOUR UNDERSTANDING OF SECTION 16.6 Suppose that speaker A in Fig. 16.23 emits a sinusoidal sound wave of frequency 500 Hz and speaker B emits a sinusoidal sound wave of frequency 1000 Hz. What sort of interference will there be between these two waves? (i) Constructive interference at various points, including point P, and destructive interference at various other points; (ii) destructive interference at various points, including point P, and constructive interference at various points; (iii) neither (i) nor (ii).

ANSWER

(iii) Constructive and destructive interference between two waves can occur only if the two waves have the same frequency. In this case the frequencies are different, so there are no points where the two waves always reinforce each other (constructive interference) or always cancel each other (destructive interference).

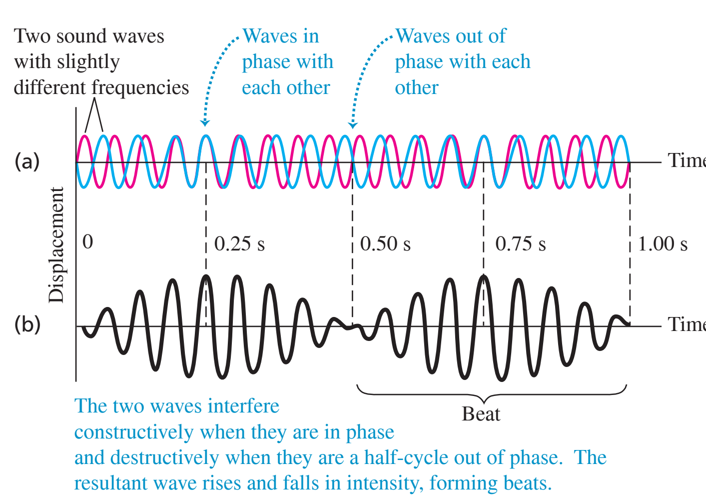
16.7 BEATS

In Section 16.6 we talked about *interference* effects that occur when two different waves with the same frequency overlap in the same region of space. Now let's look at what happens when we have two waves with equal amplitude but slightly different frequencies. This occurs, for example, when two tuning forks with slightly different frequencies are sounded together, or when two organ pipes that are supposed to have exactly the same frequency are slightly "out of tune."

Consider a particular point in space where the two waves overlap. In **Fig. 16.25a** we plot the displacements of the individual waves at this point as functions of time. The total length of the time axis represents 1 second, and the frequencies are 16 Hz (blue graph) and 18 Hz (red graph). Applying the principle of superposition, we add the two displacement functions to find the total displacement function. The result is the graph of Fig. 16.25b. At certain times the two waves are in phase; their maxima coincide and their amplitudes add. But at certain times (like $t = 0.50$ s in Fig. 16.25) the two waves are exactly *out* of phase. The two waves then cancel each other, and the total amplitude is zero.

The resultant wave in Fig. 16.25b looks like a single sinusoidal wave with an amplitude that varies from a maximum to zero and back. In this example the amplitude goes through two maxima and two minima in 1 second, so the frequency of this amplitude variation is 2 Hz. The amplitude variation causes variations of loudness called **beats,** and the frequency with which the loudness varies is called the **beat frequency.** In this example the beat frequency is the *difference* of the two frequencies. If the beat frequency is a few hertz, we hear it as a waver or pulsation in the tone.

Figure **16.25** Beats are fluctuations in amplitude produced by two sound waves of slightly different frequency, here 16 Hz and 18 Hz. (a) Individual waves. (b) Resultant wave formed by superposition of the two waves. The beat frequency is 18 Hz − 16 Hz = 2 Hz.



We can prove that the beat frequency is *always* the difference of the two frequencies f_a and f_b. Suppose f_a is larger than f_b; the corresponding periods are T_a and T_b, with $T_a < T_b$. If the two waves start out in phase at time $t = 0$, they are again in phase when the first wave has gone through exactly one more cycle than the second. This happens at a value of t equal to T_{beat}, the *period* of the beat. Let n be the number of cycles of the first wave in time T_{beat}; then the number of cycles of the second wave in the same time is $(n - 1)$, and we have the relationships

$$T_{\text{beat}} = nT_a \qquad \text{and} \qquad T_{\text{beat}} = (n - 1)T_b$$

Eliminating n between these two equations, we find

$$T_{\text{beat}} = \frac{T_a T_b}{T_b - T_a}$$

The reciprocal of the beat period is the beat *frequency*, $f_{\text{beat}} = 1/T_{\text{beat}}$, so

$$f_{\text{beat}} = \frac{T_b - T_a}{T_a T_b} = \frac{1}{T_a} - \frac{1}{T_b}$$

and finally

Beat frequency for waves a and b

$$f_{\text{beat}} = f_a - f_b \qquad (16.24)$$

Frequency of wave a

Frequency of wave b (lower than f_a)

CAUTION Beat frequency tells you only the difference in frequency between two sound waves The beat frequency for two sound waves always equals the higher frequency minus the lower frequency, so its value alone doesn't tell you which wave frequency is higher. For example, if you hear a beat frequency of 1 Hz and you know one of the sounds has frequency 256 Hz, the frequency of the other sound could be either 257 Hz or 255 Hz.

As claimed, the beat frequency is the difference of the two frequencies.

An alternative way to derive Eq. (16.24) is to write functions to describe the curves in Fig. 16.25a and then add them. Suppose that at a certain position the two waves are given by $y_a(t) = A\sin 2\pi f_a t$ and $y_b(t) = -A\sin 2\pi f_b t$. We use the trigonometric identity

$$\sin a - \sin b = 2\sin\tfrac{1}{2}(a - b)\cos\tfrac{1}{2}(a + b)$$

We can then express the total wave $y(t) = y_a(t) + y_b(t)$ as

$$y_a(t) + y_b(t) = \left[2A\sin\tfrac{1}{2}(2\pi)(f_a - f_b)t\right]\cos\tfrac{1}{2}(2\pi)(f_a + f_b)t$$

The amplitude factor (the quantity in brackets) varies slowly with frequency $\frac{1}{2}(f_a - f_b)$. The cosine factor varies with a frequency equal to the *average* frequency $\frac{1}{2}(f_a + f_b)$. The *square* of the amplitude factor, which is proportional to the intensity that the ear hears, goes through two maxima and two minima per cycle. So the beat frequency f_{beat} that is heard is twice the quantity $\frac{1}{2}(f_a - f_b)$, or just $f_a - f_b$, in agreement with Eq. (16.24).

Beats between two tones can be heard up to a beat frequency of about 6 or 7 Hz. Two piano strings or two organ pipes differing in frequency by 2 or 3 Hz sound wavery and "out of tune," although some organ stops contain two sets of pipes deliberately tuned to beat frequencies of about 1 to 2 Hz for a gently undulating effect. Listening for beats is an important technique in tuning all musical instruments. *Avoiding* beats is part of the task of flying a multiengine propeller airplane (**Fig. 16.26**).

At frequency differences greater than about 6 or 7 Hz, we no longer hear individual beats, and the sensation merges into one of *consonance* or *dissonance,* depending on the frequency ratio of the two tones. In some cases the ear perceives a tone called a *difference tone,* with a pitch equal to the beat frequency of the two tones. For example, if you listen to a whistle that produces sounds at 1800 Hz and 1900 Hz when blown, you'll hear not only these tones but also a much lower 100 Hz tone.

Figure **16.26** If the two propellers on this airplane are not precisely synchronized, the pilots, passengers, and listeners on the ground will hear beats as loud, annoying, throbbing sounds. On some airplanes the propellers are synched electronically; on others the pilot does it by ear, like tuning a piano.

TEST YOUR UNDERSTANDING OF SECTION 16.7 One tuning fork vibrates at 440 Hz, while a second tuning fork vibrates at an unknown frequency. When both tuning forks are sounded simultaneously, you hear a tone that rises and falls in intensity three times per second. What is the frequency of the second tuning fork? (i) 434 Hz; (ii) 437 Hz; (iii) 443 Hz; (iv) 446 Hz; (v) either 434 Hz or 446 Hz; (vi) either 437 Hz or 443 Hz.

ANSWER

(vi) The beat frequency is 3 Hz, so the difference between the two tuning fork frequencies is also 3 Hz. Hence the second tuning fork vibrates at a frequency of either 443 Hz or 437 Hz. You can distinguish between the two possibilities by comparing the pitches of the two tuning forks sounded one at a time: The frequency is 437 Hz if the second tuning fork has a lower pitch and 443 Hz if it has a higher pitch.

16.8 THE DOPPLER EFFECT

When a car approaches you with its horn sounding, the pitch seems to drop as the car passes. This phenomenon, first described by the 19th-century Austrian scientist Christian Doppler, is called the **Doppler effect.** When a source of sound and a listener are in motion relative to each other, the frequency of the sound heard by the listener is not the same as the source frequency. A similar effect occurs for light and radio waves; we'll return to this later in this section.

To analyze the Doppler effect for sound, we'll work out a relationship between the frequency shift and the velocities of source and listener relative to the medium (usually air) through which the sound waves propagate. To keep things simple, we consider only the special case in which the velocities of both source and listener lie along the line joining them. Let v_S and v_L be the velocity components along this line for the source and the listener, respectively, relative to the medium. We choose the positive direction for both v_S and v_L to be the direction from the listener L to the source S. The speed of sound relative to the medium, v, is always considered positive.

Moving Listener and Stationary Source

Let's think first about a listener L moving with velocity v_L toward a stationary source S (**Fig. 16.27**). The source emits a sound wave with frequency f_S and wavelength $\lambda = v/f_S$. The figure shows four wave crests, separated by equal distances λ. The wave crests approaching the moving listener have a speed of propagation *relative to the listener* of $(v + v_L)$. So the frequency f_L with which the crests arrive at the listener's position (that is, the frequency the listener hears) is

$$f_L = \frac{v + v_L}{\lambda} = \frac{v + v_L}{v/f_S} \tag{16.25}$$

or

$$f_L = \left(\frac{v + v_L}{v}\right) f_S = \left(1 + \frac{v_L}{v}\right) f_S \quad \text{(moving listener, stationary source)} \tag{16.26}$$

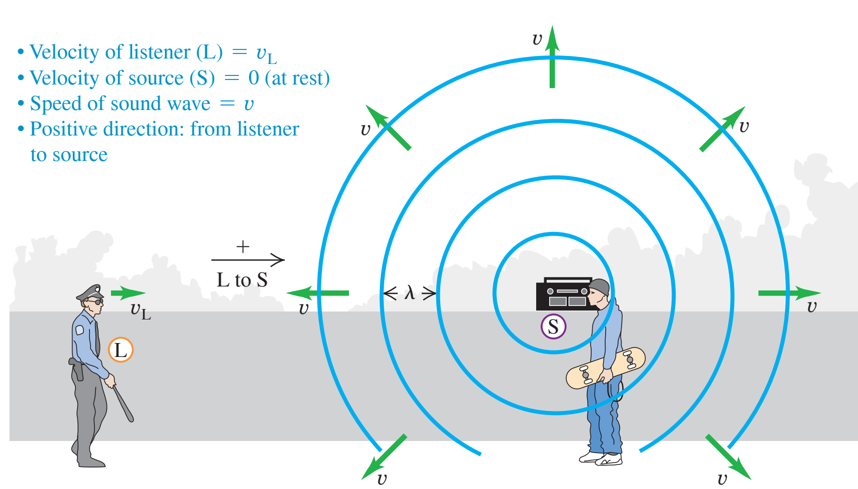


Figure 16.27 A listener moving toward a stationary source hears a frequency that is higher than the source frequency. This is because the relative speed of listener and wave is greater than the wave speed v.

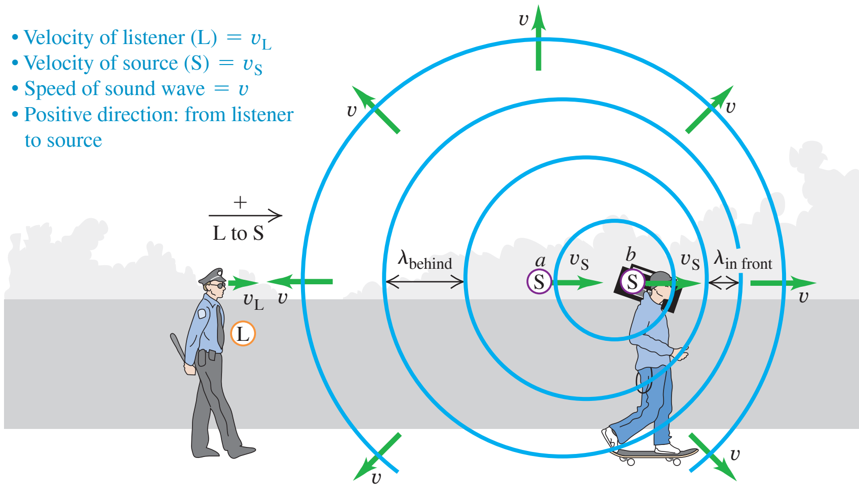


Figure **16.28** Wave crests emitted by a source moving from *a* to *b* are crowded together in front of the source (to the right of this source) and stretched out behind it (to the left of this source).

So a listener moving toward a source ($v_L > 0$), as in Fig. 16.27, hears a higher frequency (higher pitch) than does a stationary listener. A listener moving away from the source ($v_L < 0$) hears a lower frequency (lower pitch).

Moving Source and Moving Listener

Now suppose the source is also moving, with velocity v_S (**Fig. 16.28**). The wave speed relative to the wave medium (air) is still v; it is determined by the properties of the medium and is not changed by the motion of the source. But the wavelength is no longer equal to v/f_S. Here's why. The time for emission of one cycle of the wave is the period $T = 1/f_S$. During this time, the wave travels a distance $vT = v/f_S$ and the source moves a distance $v_S T = v_S/f_S$. The wavelength is the distance between successive wave crests, and this is determined by the *relative* displacement of source and wave. As Fig. 16.28 shows, this is different in front of and behind the source. In the region to the right of the source in Fig. 16.28 (that is, in front of the source), the wavelength is

$$\lambda_{\text{in front}} = \frac{v}{f_S} - \frac{v_S}{f_S} = \frac{v - v_S}{f_S} \qquad \text{(wavelength in front of a moving source)} \qquad (16.27)$$

In the region to the left of the source (that is, behind the source), it is

$$\lambda_{\text{behind}} = \frac{v + v_S}{f_S} \qquad \text{(wavelength behind a moving source)} \qquad (16.28)$$

The waves in front of and behind the source are compressed and stretched out, respectively, by the motion of the source.

To find the frequency heard by the listener behind the source, we substitute Eq. (16.28) into the first form of Eq. (16.25):

$$f_L = \frac{v + v_L}{\lambda_{\text{behind}}} = \frac{v + v_L}{(v + v_S)/f_S}$$

Doppler effect for moving listener L and moving source S:

Frequency heard by listener; Velocity of listener (+ if from L toward S, − if opposite)

$$f_L = \frac{v + v_L}{v + v_S} f_S \qquad (16.29)$$

Frequency emitted by source; Speed of sound; Velocity of source (+ if from L toward S, − if opposite)

Figure 16.29 The Doppler effect explains why the siren on a fire engine or ambulance has a high pitch ($f_L > f_S$) when it is approaching you ($v_S < 0$) and a low pitch ($f_L < f_S$) when it is moving away ($v_S > 0$).

Although we derived it for the particular situation shown in Fig. 16.28, Eq. (16.29) includes *all* possibilities for motion of source and listener (relative to the medium) along the line joining them. If the listener happens to be at rest in the medium, v_L is zero. When both source and listener are at rest or have the same velocity relative to the medium, $v_L = v_S$ and $f_L = f_S$. Whenever the direction of the source or listener velocity is opposite to the direction from the listener toward the source (which we have defined as positive), the corresponding velocity to be used in Eq. (16.29) is negative.

As an example, the frequency heard by a listener at rest ($v_L = 0$) is $f_L = [v/(v + v_S)]f_S$. If the source is moving toward the listener (in the negative direction), then $v_S < 0$, $f_L > f_S$, and the listener hears a higher frequency than that emitted by the source. If instead the source is moving away from the listener (in the positive direction), then $v_S > 0$, $f_L < f_S$, and the listener hears a lower frequency. This explains the change in pitch that you hear from the siren of an ambulance as it passes you (**Fig. 16.29**).

PROBLEM-SOLVING STRATEGY 16.2 Doppler Effect

IDENTIFY *the relevant concepts:* The Doppler effect occurs whenever the source of waves, the wave detector (listener), or both are in motion.

SET UP *the problem* using the following steps:

1. Establish a coordinate system, with the positive direction from the listener toward the source. Carefully determine the signs of all relevant velocities. A velocity in the direction from the listener toward the source is positive; a velocity in the opposite direction is negative. All velocities must be measured relative to the air in which the sound travels.
2. Use consistent subscripts to identify the various quantities: S for source and L for listener.
3. Identify which unknown quantities are the target variables.

EXECUTE *the solution* as follows:

1. Use Eq. (16.29) to relate the frequencies at the source and the listener, the sound speed, and the velocities of the source and the listener according to the sign convention of step 1. If the source is moving, you can find the wavelength measured by the listener using Eq. (16.27) or (16.28).
2. When a wave is reflected from a stationary or moving surface, solve the problem in two steps. In the first, the surface is the "listener"; the frequency with which the wave crests arrive at the surface is f_L. In the second, the surface is the "source," emitting waves with this same frequency f_L. Finally, determine the frequency heard by a listener detecting this new wave.

EVALUATE *your answer:* Is the *direction* of the frequency shift reasonable? If the source and the listener are moving toward each other, $f_L > f_S$; if they are moving apart, $f_L < f_S$. If the source and the listener have no relative motion, $f_L = f_S$.

EXAMPLE 16.14 Doppler effect I: Wavelengths

WITH VARIATION PROBLEMS

A police car's siren emits a sinusoidal wave with frequency $f_S = 300$ Hz. The speed of sound is 340 m/s and the air is still. (a) Find the wavelength of the waves if the siren is at rest. (b) Find the wavelengths of the waves in front of and behind the siren if it is moving at 30 m/s.

IDENTIFY and SET UP In part (a) there is no Doppler effect because neither source nor listener is moving with respect to the air; $v = \lambda f$ gives the wavelength. **Figure 16.30** shows the situation in part (b): The source is in motion, so we find the wavelengths using Eqs. (16.27) and (16.28) for the Doppler effect.

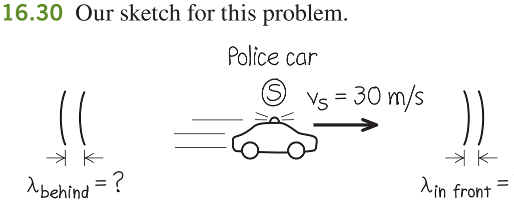
EXECUTE (a) When the source is at rest,

$$\lambda = \frac{v}{f_S} = \frac{340 \text{ m/s}}{300 \text{ Hz}} = 1.13 \text{ m}$$

(b) From Eq. (16.27), in front of the siren

$$\lambda_{\text{in front}} = \frac{v - v_S}{f_S} = \frac{340 \text{ m/s} - 30 \text{ m/s}}{300 \text{ Hz}} = 1.03 \text{ m}$$

Figure 16.30 Our sketch for this problem.



From Eq. (16.28), behind the siren

$$\lambda_{\text{behind}} = \frac{v + v_S}{f_S} = \frac{340 \text{ m/s} + 30 \text{ m/s}}{300 \text{ Hz}} = 1.23 \text{ m}$$

EVALUATE The wavelength is shorter in front of the siren and longer behind it, as we expect.

KEYCONCEPT If a source of sound is moving through still air, a listener behind the source hears a sound of increased wavelength. A listener in front of the source hears a sound of decreased wavelength.

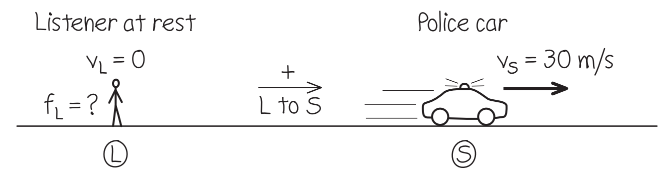
EXAMPLE 16.15 Doppler effect II: Frequencies

WITH VARIATION PROBLEMS

If a listener L is at rest and the siren in Example 16.14 is moving away from L at 30 m/s, what frequency does the listener hear?

IDENTIFY and SET UP Our target variable is the frequency f_L heard by a listener behind the moving source. **Figure 16.31** shows the situation. We have $v_L = 0$ and $v_S = +30$ m/s (positive, since the velocity of the source is in the direction from listener to source).

Figure **16.31** Our sketch for this problem.



EXECUTE From Eq. (16.29),

$$f_L = \frac{v}{v + v_S} f_S = \frac{340 \text{ m/s}}{340 \text{ m/s} + 30 \text{ m/s}}(300 \text{ Hz}) = 276 \text{ Hz}$$

EVALUATE The source and listener are moving apart, so $f_L < f_S$. Here's a check on our numerical result. From Example 16.14, the wavelength behind the source (where the listener in Fig. 16.31 is located) is 1.23 m. The wave speed relative to the stationary listener is $v = 340$ m/s even though the source is moving, so

$$f_L = \frac{v}{\lambda} = \frac{340 \text{ m/s}}{1.23 \text{ m}} = 276 \text{ Hz}$$

KEYCONCEPT If a source of sound is moving through still air, a listener behind the source hears a sound of decreased frequency. A listener in front of the source hears a sound of increased frequency.

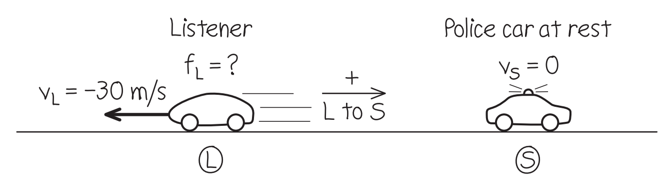
EXAMPLE 16.16 Doppler effect III: A moving listener

WITH VARIATION PROBLEMS

If the siren is at rest and the listener is moving away from it at 30 m/s, what frequency does the listener hear?

IDENTIFY and SET UP Again our target variable is f_L, but now L is in motion and S is at rest. **Figure 16.32** shows the situation. The velocity of the listener is $v_L = -30$ m/s (negative, since the motion is in the direction from source to listener).

Figure **16.32** Our sketch for this problem.



EXECUTE From Eq. (16.29),

$$f_L = \frac{v + v_L}{v} f_S = \frac{340 \text{ m/s} + (-30 \text{ m/s})}{340 \text{ m/s}}(300 \text{ Hz}) = 274 \text{ Hz}$$

EVALUATE Again the source and listener are moving apart, so $f_L < f_S$. Note that the *relative velocity* of source and listener is the same as in Example 16.15, but the Doppler shift is different because v_S and v_L are different.

KEYCONCEPT If a listener is moving away from a source of sound, the listener hears a sound of decreased frequency. If the listener is moving toward the source, the listener hears a sound of increased frequency.

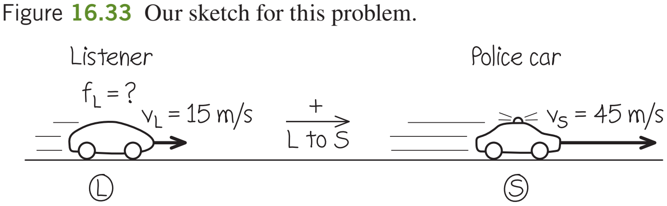
EXAMPLE 16.17 Doppler effect IV: Moving source, moving listener

WITH VARIATION PROBLEMS

The siren is moving away from the listener with a speed of 45 m/s relative to the air, and the listener is moving toward the siren with a speed of 15 m/s relative to the air. What frequency does the listener hear?

IDENTIFY and SET UP Now *both* L and S are in motion (**Fig. 16.33**). Again our target variable is f_L. Both the source velocity $v_S = +45$ m/s and the listener's velocity $v_L = +15$ m/s are positive because both velocities are in the direction from listener to source.

Figure **16.33** Our sketch for this problem.



EXECUTE From Eq. (16.29),

$$f_L = \frac{v + v_L}{v + v_S} f_S = \frac{340 \text{ m/s} + 15 \text{ m/s}}{340 \text{ m/s} + 45 \text{ m/s}}(300 \text{ Hz}) = 277 \text{ Hz}$$

EVALUATE As in Examples 16.15 and 16.16, the source and listener again move away from each other at 30 m/s, so again $f_L < f_S$. But f_L is different in all three cases because the Doppler effect for sound depends on how the source and listener are moving relative to the *air*, not simply on how they move relative to each other.

KEYCONCEPT When a listener and source of sound are both moving relative to the air, the listener hears a sound of decreased frequency if the listener and source are moving apart. The listener hears a sound of increased frequency if the listener and source are moving closer together.

EXAMPLE 16.18 Doppler effect V: A double Doppler shift

WITH VARIATION PROBLEMS

The police car is moving toward a warehouse at 30 m/s. What frequency does the driver hear reflected from the warehouse?

IDENTIFY This situation has *two* Doppler shifts (**Fig. 16.34**). In the first shift, the warehouse is the stationary "listener." The frequency of sound reaching the warehouse, which we call f_W, is greater than 300 Hz because the source is approaching. In the second shift, the warehouse acts as a source of sound with frequency f_W, and the listener is the driver of the police car; she hears a frequency greater than f_W because she is approaching the source.

SET UP To determine f_W, we use Eq. (16.29) with f_L replaced by f_W. For this part of the problem, $v_L = v_W = 0$ (the warehouse is at rest) and $v_S = -30$ m/s (the siren is moving in the negative direction from source to listener).

To determine the frequency heard by the driver (our target variable), we again use Eq. (16.29) but now with f_S replaced by f_W. For this second part of the problem, $v_S = 0$ because the stationary warehouse is the source and the velocity of the listener (the driver) is $v_L = +30$ m/s. (The listener's velocity is positive because it is in the direction from listener to source.)

EXECUTE The frequency reaching the warehouse is

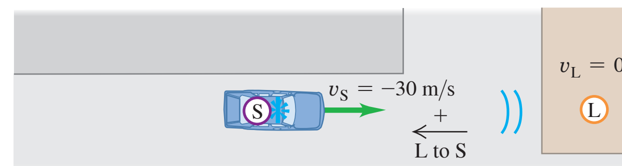
$$f_W = \frac{v}{v + v_S} f_S = \frac{340 \text{ m/s}}{340 \text{ m/s} + (-30 \text{ m/s})}(300 \text{ Hz})$$
$$= 329 \text{ Hz}$$

Then the frequency heard by the driver is

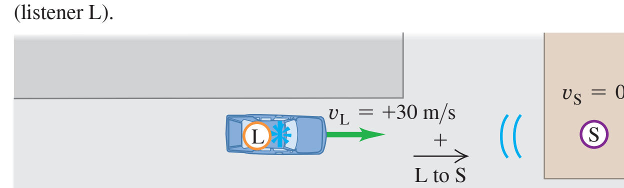
$$f_L = \frac{v + v_L}{v} f_W = \frac{340 \text{ m/s} + 30 \text{ m/s}}{340 \text{ m/s}}(329 \text{ Hz}) = 358 \text{ Hz}$$

Figure 16.34 Two stages of the sound wave's motion from the police car to the warehouse and back to the police car.

(a) Sound travels from police car's siren (source S) to warehouse ("listener" L).



(b) Reflected sound travels from warehouse (source S) to police car (listener L).



EVALUATE Because there are two Doppler shifts, the reflected sound heard by the driver has an even higher frequency than the sound heard by a stationary listener in the warehouse.

KEYCONCEPT If a source of sound moves relative to a wall (or other reflecting surface), there are two shifts in the frequency of the sound: The frequency received by and reflected from the wall is shifted compared to the sound emitted by the source, and the frequency received back at the source is shifted compared to the sound reflected from the wall. Both shifts increase the frequency if the source is approaching the wall, and both shifts decrease the frequency if the source is moving away from the wall.

Doppler Effect for Electromagnetic Waves

In the Doppler effect for sound, the velocities v_L and v_S are always measured relative to the *air* or whatever medium we are considering. There is also a Doppler effect for *electromagnetic* waves in empty space, such as light waves or radio waves. In this case there is no medium that we can use as a reference to measure velocities, and all that matters is the *relative* velocity of source and receiver. (By contrast, the Doppler effect for sound does not depend simply on this relative velocity, as discussed in Example 16.17.)

To derive the expression for the Doppler frequency shift for light, we have to use the special theory of relativity. We'll discuss this in Chapter 37, but for now we quote the result without derivation. The wave speed is the speed of light, usually denoted by c, and it is the same for both source and receiver. In the frame of reference in which the receiver is at rest, the source is moving away from the receiver with velocity v. (If the source is *approaching* the receiver, v is negative.) The source frequency is again f_S. The frequency f_R measured by the receiver R (the frequency of arrival of the waves at the receiver) is then

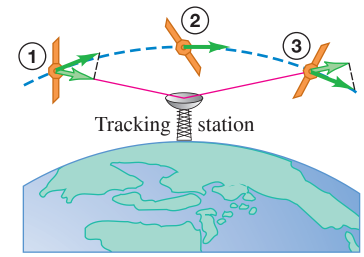
$$f_R = \sqrt{\frac{c - v}{c + v}} f_S \qquad \text{(Doppler effect for light)} \qquad (16.30)$$

When v is positive, the source is moving directly *away* from the receiver and f_R is always *less* than f_S; when v is negative, the source is moving directly *toward* the receiver and f_R is *greater* than f_S. The qualitative effect is the same as for sound, but the quantitative relationship is different.

A familiar application of the Doppler effect for radio waves is the radar device mounted on the side window of a police car to check other cars' speeds. The electromagnetic wave emitted by the device is reflected from a moving car, which acts as a moving source, and the wave reflected back to the device is Doppler-shifted in frequency. The transmitted and reflected signals are combined to produce beats, and the speed can be computed from the frequency of the beats. Similar techniques ("Doppler radar") are used to measure wind velocities in the atmosphere.

The Doppler effect is also used to track satellites and other space vehicles. In **Fig. 16.35** a satellite emits a radio signal with constant frequency f_S. As the satellite orbits past, it first approaches and then moves away from the receiver; the frequency f_R of the signal received on earth changes from a value greater than f_S to a value less than f_S as the satellite passes overhead.

Figure **16.35** Change of velocity component along the line of sight of a satellite passing a tracking station. The frequency received at the tracking station changes from high to low as the satellite passes overhead.



TEST YOUR UNDERSTANDING OF SECTION 16.8 You are at an outdoor concert with a wind blowing at 10 m/s from the performers toward you. Is the sound you hear Doppler-shifted? If so, is it shifted to lower or higher frequencies?

ANSWER

no The air (the medium for sound waves) is moving from the source toward the listener. Hence, relative to the air, both the source and the listener are moving in the direction from listener to source. So both velocities are positive and $v_S = v_L = +10$ m/s. The equality of these two velocities means that the numerator and the denominator in Eq. (16.29) are the same, so $f_L = f_S$ and there is *no* Doppler shift.

16.9 SHOCK WAVES

You may have experienced "sonic booms" caused by an airplane flying overhead faster than the speed of sound. We can see qualitatively why this happens from **Fig. 16.36**. Let v_S denote the *speed* of the airplane relative to the air, so that it is always positive. The motion of the airplane through the air produces sound; if v_S is less than the speed of sound v, the waves in front of the airplane are crowded together with a wavelength given by Eq. (16.27):

$$\lambda_{\text{in front}} = \frac{v - v_S}{f_S}$$

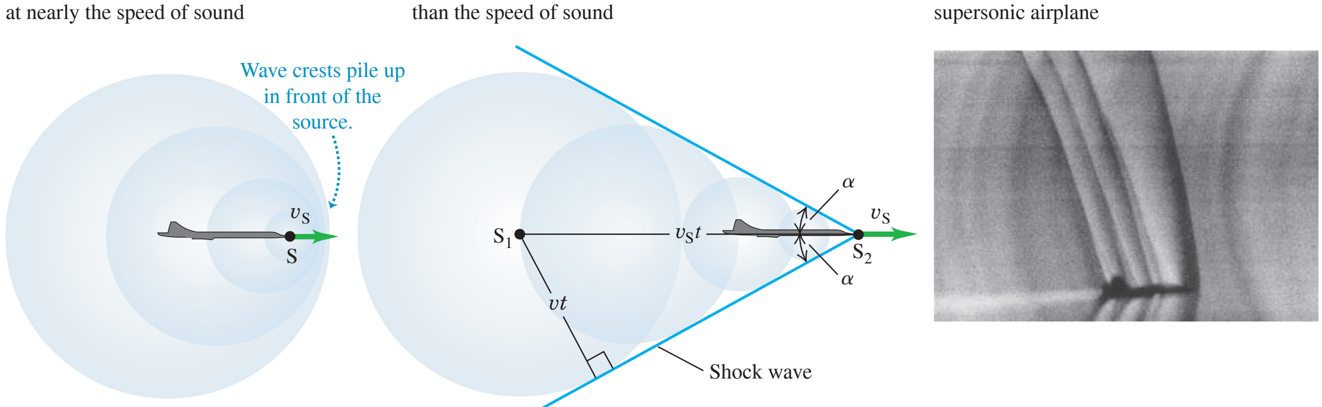
As the speed v_S of the airplane approaches the speed of sound v, the wavelength approaches zero and the wave crests pile up on each other (Fig. 16.36a). The airplane must exert a large force to compress the air in front of it; by Newton's third law, the air exerts an equally large force back on the airplane. Hence there is a large increase in aerodynamic drag (air resistance) as the airplane approaches the speed of sound, a phenomenon known as the "sound barrier."

Figure **16.36** Wave crests around a sound source S moving (a) slightly slower than the speed of sound v and (b) faster than the sound speed v. (c) This photograph shows a T-38 jet airplane moving at 1.1 times the speed of sound. Separate shock waves are produced by the nose, wings, and tail. The angles of these waves vary because the air speeds up and slows down as it moves around the airplane, so the relative speed v_S of the airplane and air is different for shock waves produced at different points.

(a) Sound source S (airplane) moving at nearly the speed of sound

(b) Sound source moving faster than the speed of sound

(c) Shock waves around a supersonic airplane



When v_S is greater in magnitude than v, the source of sound is **supersonic,** and Eqs. (16.27) and (16.29) for the Doppler effect no longer describe the sound wave in front of the source. Figure 16.36b shows a cross section of what happens. As the airplane moves, it displaces the surrounding air and produces sound. A series of wave crests is emitted from the nose of the airplane; each spreads out in a circle centered at the position of the airplane when it emitted the crest. After a time t the crest emitted from point S_1 has spread to a circle with radius vt, and the airplane has moved a greater distance $v_S t$ to position S_2. You can see that the circular crests interfere constructively at points along the blue line that makes an angle α with the direction of the airplane velocity, leading to a very-large-amplitude wave crest along this line. This large-amplitude crest is called a **shock wave** (Fig. 16.36c).

From the right triangle in Fig. 16.36b we can see that $\sin\alpha = vt/v_S t$, or

Shock wave produced by sound source moving faster than sound:

$$\underbrace{\sin\alpha}_{\text{Angle of shock wave}} = \frac{v \;\leftarrow \text{Speed of sound}}{v_S \;\leftarrow \text{Speed of source}} \tag{16.31}$$

Figure 16.37 The first supersonic airplane, the Bell X-1, was shaped much like a 50 caliber bullet—which was known to be able to travel faster than sound.

The ratio v_S/v is called the **Mach number.** It is greater than unity for all supersonic speeds, and $\sin\alpha$ in Eq. (16.31) is the reciprocal of the Mach number. The first person to break the sound barrier was Capt. Chuck Yeager of the U.S. Air Force, flying the Bell X-1 at Mach 1.06 on October 14, 1947 (**Fig. 16.37**).

Shock waves are actually three-dimensional; a shock wave forms a *cone* around the direction of motion of the source. If the source (possibly a supersonic jet airplane or a rifle bullet) moves with constant velocity, the angle α is constant, and the shock-wave cone moves along with the source. It's the arrival of this shock wave that causes the sonic boom you hear after a supersonic airplane has passed by. In front of the shock-wave cone, there is no sound. Inside the cone a stationary listener hears the Doppler-shifted sound of the airplane moving away.

CAUTION Shock waves A shock wave is produced *continuously* by any object that moves through the air at supersonic speed, not only at the instant that it "breaks the sound barrier." The sound waves that combine to form the shock wave, as in Fig. 16.36b, are created by the motion of the object itself, not by any sound source that the object may carry. The cracking noises of a bullet and of the tip of a circus whip are due to their supersonic motion. A supersonic jet airplane may have very loud engines, but these do not cause the shock wave. If the pilot were to shut the engines off, the airplane would continue to produce a shock wave as long as its speed remained supersonic.

Shock waves have applications outside of aviation. They are used to break up kidney stones and gallstones without invasive surgery, using a technique with the impressive name *extracorporeal shock-wave lithotripsy.* A shock wave produced outside the body is focused by a reflector or acoustic lens so that as much of it as possible converges on the stone. When the resulting stresses in the stone exceed its tensile strength, it breaks into small pieces and can be eliminated. This technique requires accurate determination of the location of the stone, which may be done using ultrasonic imaging techniques (see Fig. 16.9).

EXAMPLE 16.19 Sonic boom from a supersonic airplane

An airplane is flying at Mach 1.75 at an altitude of 8000 m, where the speed of sound is 320 m/s. How long after the plane passes directly overhead will you hear the sonic boom?

IDENTIFY and SET UP The shock wave forms a cone trailing backward from the airplane, so the problem is really asking for how much time elapses from when the airplane flies overhead to when the shock wave reaches you at point L (**Fig. 16.38**). During the time t (our target variable) since the airplane traveling at speed v_S passed overhead, it has traveled a distance $v_S t$. Equation (16.31) gives the shock cone angle α; we use trigonometry to solve for t.

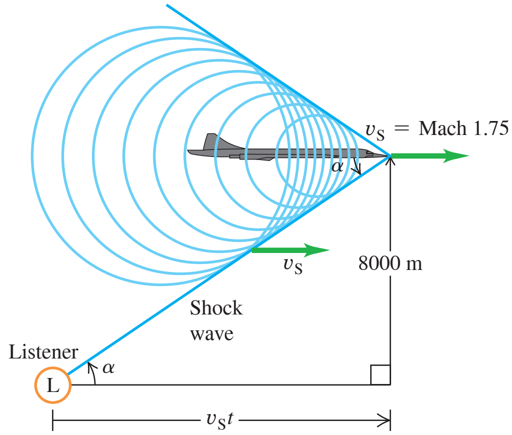
EXECUTE From Eq. (16.31) the angle α of the shock cone is

$$\alpha = \arcsin\frac{1}{1.75} = 34.8^\circ$$

The speed of the plane is the speed of sound multiplied by the Mach number:

$$v_S = (1.75)(320\ \text{m/s}) = 560\ \text{m/s}$$

Figure **16.38** You hear a sonic boom when the shock wave reaches you at L (*not* just when the plane breaks the sound barrier). A listener to the right of L has not yet heard the sonic boom but will shortly; a listener to the left of L has already heard the sonic boom.



From Fig. 16.38 we have

$$\tan\alpha = \frac{8000\ \text{m}}{v_S t}$$

$$t = \frac{8000\ \text{m}}{(560\ \text{m/s})(\tan 34.8°)} = 20.5\ \text{s}$$

EVALUATE You hear the boom 20.5 s after the airplane passes overhead, at which time it has traveled $(560\ \text{m/s})(20.5\ \text{s}) = 11.5\ \text{km}$ since it passed overhead. We have assumed that the speed of sound is the same at all altitudes, so that $\alpha = \arcsin v/v_S$ is constant and the shock wave forms a perfect cone. In fact, the speed of sound decreases with increasing altitude. How would this affect the value of t?

KEYCONCEPT An object moving through the air faster than the speed of sound continuously produces a cone-shaped shock wave. The angle of the cone depends on the object's Mach number (the ratio of its speed to the speed of sound).

TEST YOUR UNDERSTANDING OF SECTION 16.9 What would you hear if you were directly behind (to the left of) the supersonic airplane in Fig. 16.38? (i) A sonic boom; (ii) the sound of the airplane, Doppler-shifted to higher frequencies; (iii) the sound of the airplane, Doppler-shifted to lower frequencies; (iv) nothing.

ANSWER **(iii)** Figure 16.38 shows that there are sound waves inside the cone of the shock wave. Behind the airplane the wave crests are spread apart, just as they are behind the moving source in Fig. 16.28. Hence the waves that reach you have an increased wavelength and a lower frequency.

CHAPTER 16 SUMMARY

Sound waves: Sound consists of longitudinal waves in a medium. A sinusoidal sound wave is characterized by its frequency f and wavelength λ (or angular frequency ω and wave number k) and by its displacement amplitude A. The pressure amplitude p_{max} is directly proportional to the displacement amplitude, the wave number, and the bulk modulus B of the wave medium. (See Examples 16.1 and 16.2.)

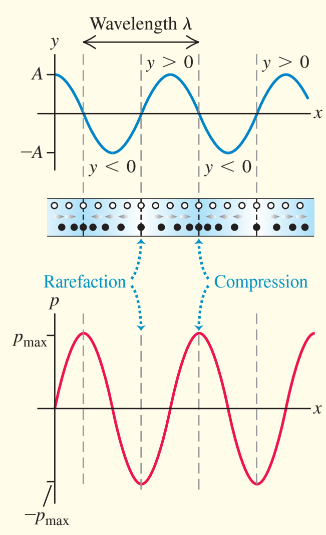
The speed of a sound wave in a fluid depends on the bulk modulus B and density ρ. If the fluid is an ideal gas, the speed can be expressed in terms of the temperature T, molar mass M, and ratio of heat capacities γ of the gas. The speed of longitudinal waves in a solid rod depends on the density and Young's modulus Y. (See Examples 16.3 and 16.4.)

$$p_{max} = BkA \qquad (16.5)$$
(sinusoidal sound wave)

$$v = \sqrt{\frac{B}{\rho}} \qquad (16.7)$$
(longitudinal wave in a fluid)

$$v = \sqrt{\frac{\gamma RT}{M}} \qquad (16.10)$$
(sound wave in an ideal gas)

$$v = \sqrt{\frac{Y}{\rho}} \qquad (16.8)$$
(longitudinal wave in a solid rod)



Intensity and sound intensity level: The intensity I of a sound wave is the time average rate at which energy is transported by the wave, per unit area. For a sinusoidal wave, the intensity can be expressed in terms of the displacement amplitude A or the pressure amplitude $p_{\max}$. (See Examples 16.5–16.7.)

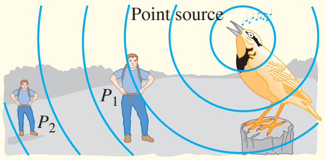
The sound intensity level β of a sound wave is a logarithmic measure of its intensity. It is measured relative to I_0, an arbitrary intensity defined to be $10^{-12}\ \mathrm{W/m^2}$. Sound intensity levels are expressed in decibels (dB). (See Examples 16.8 and 16.9.)

$$I = \tfrac{1}{2}\sqrt{\rho B}\,\omega^2 A^2 = \frac{p_{\max}{}^2}{2\rho v} = \frac{p_{\max}{}^2}{2\sqrt{\rho B}} \qquad (16.12),\ (16.14)$$

(intensity of a sinusoidal sound wave in a fluid)

$$\beta = (10\ \mathrm{dB})\log\frac{I}{I_0} \qquad (16.15)$$

(definition of sound intensity level)



Standing sound waves: Standing sound waves can be set up in a pipe or tube. A closed end is a displacement node and a pressure antinode; an open end is a displacement antinode and a pressure node. For a pipe of length L open at both ends, the normal-mode frequencies are integer multiples of the sound speed divided by $2L$. For a stopped pipe (one that is open at only one end), the normal-mode frequencies are the odd multiples of the sound speed divided by $4L$. (See Examples 16.10 and 16.11.)

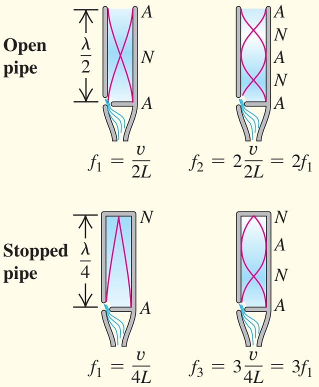
A pipe or other system with normal-mode frequencies can be driven to oscillate at any frequency. A maximum response, or resonance, occurs if the driving frequency is close to one of the normal-mode frequencies of the system. (See Example 16.12.)

$$f_n = \frac{nv}{2L} \quad (n = 1, 2, 3, \ldots) \qquad (16.18)$$

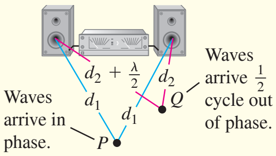
(open pipe)

$$f_n = \frac{nv}{4L} \quad (n = 1, 3, 5, \ldots) \qquad (16.22)$$

(stopped pipe)



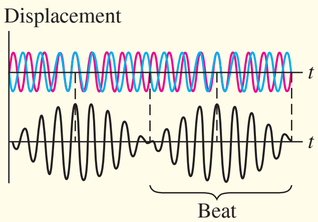
Interference: When two or more waves overlap in the same region of space, the resulting effects are called interference. The resulting amplitude can be either larger or smaller than the amplitude of each individual wave, depending on whether the waves are in phase (constructive interference) or out of phase (destructive interference). (See Example 16.13.)



Beats: Beats are heard when two tones with slightly different frequencies f_a and f_b are sounded together. The beat frequency f_{beat} is the difference between f_a and f_b.

$$f_{\text{beat}} = f_a - f_b \qquad (16.24)$$

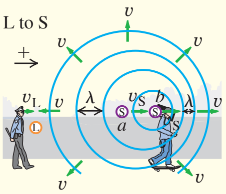
(beat frequency)



Doppler effect: The Doppler effect for sound is the frequency shift that occurs when there is motion of a source of sound, a listener, or both, relative to the medium. The source and listener frequencies f_S and f_L are related by the source and listener velocities v_S and v_L relative to the medium and to the speed of sound v. (See Examples 16.14–16.18.)

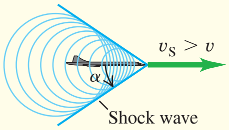
$$f_L = \frac{v + v_L}{v + v_S} f_S \qquad (16.29)$$

(Doppler effect, moving source and moving listener)



Shock waves: A sound source moving with a speed v_S greater than the speed of sound v creates a shock wave. The wave front is a cone with angle α. (See Example 16.19.)

$$\sin\alpha = \frac{v}{v_S} \quad \text{(shock wave)} \qquad (16.31)$$



Chapter 16 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 16.5, 16.6, 16.7, 16.8, and 16.9 (Section 16.3) before attempting these problems.

VP16.9.1 A 256 Hz sound wave in air (density 1.20 kg/m^3, speed of sound 344 m/s) has intensity 5.50×10^{-8} W/m^2. (a) What is the wave's pressure amplitude? (b) If the intensity remains the same but the frequency is doubled to 512 Hz, how does this affect the pressure amplitude?

VP16.9.2 At a certain distance from a fire alarm, the sound intensity level is 85.0 dB. (a) What is the intensity of this sound? (b) How many times greater is the intensity of this sound than that of a 67.0 dB sound?

VP16.9.3 A lion can produce a roar with a sound intensity level of 114 dB at a distance of 1.00 m. What is the sound intensity level at a distance of (a) 4.00 m and (b) 15.8 m from the lion? Assume that intensity obeys the inverse-square law.

VP16.9.4 The sound intensity level inside a typical modern airliner in flight is 66.0 dB. The air in the cabin has density 0.920 kg/m^3 (less than in the atmosphere at sea level) and speed of sound 344 m/s. (a) What is the pressure amplitude of this sound? (b) If the pressure amplitude were increased by a factor of 10.0, what would the new sound intensity level be?

Be sure to review EXAMPLES 16.11 and 16.12 (Sections 16.4 and 16.5) before attempting these problems.

VP16.12.1 A particular open organ pipe has a fundamental frequency of 220 Hz (known to musicians as A_3 or "A below middle C") when the speed of sound waves in air is 344 m/s. (a) What is the length of this pipe? (b) The third harmonic of this pipe has the same frequency as the fundamental frequency of a stopped pipe. What is the length of this stopped pipe?

VP16.12.2 You have two organ pipes, one open and one stopped. Which harmonic (if any) of the stopped pipe has the same frequency as the third harmonic of the open pipe if the stopped pipe length is (a) $\frac{1}{6}$, (b) $\frac{1}{2}$, or (c) $\frac{1}{3}$ that of the open pipe?

VP16.12.3 One of the strings of a bass viol is 0.680 m long and has a fundamental frequency of 165 Hz. (a) What is the speed of waves on this string? (b) When this string vibrates at its fundamental frequency, it causes the air in a nearby stopped organ pipe to vibrate at that pipe's fundamental frequency. The speed of sound in the pipe is 344 m/s. What is the length of this pipe?

VP16.12.4 A stopped pipe 1.00 m in length is filled with helium at 20°C (speed of sound 999 m/s). When the helium in this pipe vibrates at its third harmonic frequency, it causes the air at 20°C (speed of sound 344 m/s) in a nearby open pipe to vibrate at its fifth harmonic frequency. What are the frequency and wavelength of the sound wave (a) in the helium in the stopped pipe and (b) in the air in the open pipe? (c) What is the length of the open pipe?

Be sure to review EXAMPLES 16.14, 16.15, 16.16, 16.17, and 16.18 (Section 16.8) before attempting these problems.

VP16.18.1 The siren on an ambulance emits a sound of frequency 2.80×10^3 Hz. If the ambulance is traveling at 26.0 m/s (93.6 km/h, or 58.2 mi/h), the speed of sound is 340 m/s, and the air is still, what are the frequency and wavelength that you hear if you are standing (a) in front of the ambulance or (b) behind the ambulance?

VP16.18.2 A stationary bagpiper is playing a Highland bagpipe, in which one reed produces a continuous sound of frequency 440 Hz. The air is still and the speed of sound is 340 m/s. (a) What is the wavelength of the sound wave produced by the bagpipe? What are the frequency and wavelength of the sound wave that a bicyclist hears if she is (b) approaching the bagpiper at 10.0 m/s or (c) moving away from the bagpiper at 10.0 m/s?

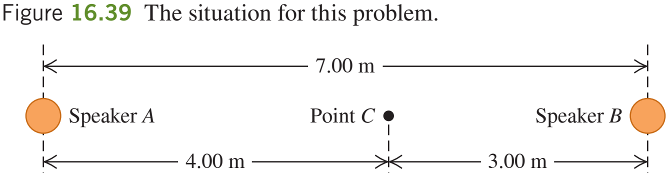
VP16.18.3 A police car moving east at 40.0 m/s is chasing a speeding sports car moving east at 35.0 m/s. The police car's siren has frequency 1.20×10^3 Hz, the speed of sound is 340 m/s, and the air is still. (a) What is the frequency of sound that the driver of the speeding sports car hears? (b) If the speeding sports car were to turn around and drive west at 35.0 m/s toward the approaching police car, what frequency would the driver of the sports car hear?

VP16.18.4 For a scene in an action movie, a car drives at 25.0 m/s directly toward a wall. The car's horn is on continuously and produces a sound of frequency 415 Hz. (a) If the speed of sound is 340 m/s and the air is still, what is the frequency of the sound that the driver of the car hears reflected from the wall? (b) How fast would the car have to move for the reflected sound that the driver hears to have frequency 495 Hz?

BRIDGING PROBLEM Loudspeaker Interference

Loudspeakers *A* and *B* are 7.00 m apart and vibrate in phase at 172 Hz. They radiate sound uniformly in all directions. Their acoustic power outputs are 8.00×10^{-4} W and 6.00×10^{-5} W, respectively. The air temperature is 20°C. (a) Determine the difference in phase of the two signals at a point *C* along the line joining *A* and *B*, 3.00 m from *B* and 4.00 m from *A* (**Fig. 16.39**). (b) Determine the intensity and sound intensity level at *C* from speaker *A* alone (with *B* turned off) and from speaker *B* alone (with *A* turned off). (c) Determine the intensity and sound intensity level at *C* from both speakers together.

Figure **16.39** The situation for this problem.



SOLUTION GUIDE

IDENTIFY and SET UP

1. Choose the equations that relate power, distance from the source, intensity, pressure amplitude, and sound intensity level.
2. Decide how you'll determine the phase difference in part (a). Once you have found the phase difference, how can you use it to find the amplitude of the combined wave at *C* due to both sources?
3. List the unknown quantities for each part of the problem and identify your target variables.

Continued

EXECUTE

4. Determine the phase difference at point C.
5. Find the intensity, sound intensity level, and pressure amplitude at C due to each speaker alone.
6. Use your results from steps 4 and 5 to find the pressure amplitude at C due to both loudspeakers together.
7. Use your result from step 6 to find the intensity and sound intensity level at C due to both loudspeakers together.

EVALUATE

8. How do your results from part (c) for intensity and sound intensity level at C compare to those from part (b)? Does this make sense?
9. What result would you have gotten in part (c) if you had (incorrectly) combined the *intensities* from A and B directly, rather than (correctly) combining the *pressure amplitudes* as you did in step 6?

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/ or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q16.1 When sound travels from air into water, does the frequency of the wave change? The speed? The wavelength? Explain your reasoning.

Q16.2 The hero of a western movie listens for an oncoming train by putting his ear to the track. Why does this method give an earlier warning of the approach of a train than just listening in the usual way?

Q16.3 Would you expect the pitch (or frequency) of an organ pipe to increase or decrease with increasing temperature? Explain.

Q16.4 In most modern wind instruments the pitch is changed by using keys or valves to change the length of the vibrating air column. The bugle, however, has no valves or keys, yet it can play many notes. How might this be possible? Are there restrictions on what notes a bugle can play?

Q16.5 Symphonic musicians always "warm up" their wind instruments by blowing into them before a performance. What purpose does this serve?

Q16.6 In a popular and amusing science demonstration, a person inhales helium and then his voice becomes high and squeaky. Why does this happen? (*Warning:* Inhaling too much helium can cause unconsciousness or death.)

Q16.7 Lane dividers on highways sometimes have regularly spaced ridges or ripples. When the tires of a moving car roll along such a divider, a musical note is produced. Why? Explain how this phenomenon could be used to measure the car's speed.

Q16.8 (a) Does a sound level of 0 dB mean that there is no sound? (b) Is there any physical meaning to a sound having a negative intensity level? If so, what is it? (c) Does a sound intensity of zero mean that there is no sound? (d) Is there any physical meaning to a sound having a negative intensity? Why?

Q16.9 Which has a more direct influence on the loudness of a sound wave: the *displacement* amplitude or the *pressure* amplitude? Explain.

Q16.10 If the pressure amplitude of a sound wave is halved, by what factor does the intensity of the wave decrease? By what factor must the pressure amplitude of a sound wave be increased in order to increase the intensity by a factor of 16? Explain.

Q16.11 Does the sound intensity level β obey the inverse-square law? Why?

Q16.12 A small fraction of the energy in a sound wave is absorbed by the air through which the sound passes. How does this modify the inverse-square relationship between intensity and distance from the source? Explain.

Q16.13 A small metal band is slipped onto one of the tines of a tuning fork. As this band is moved closer and closer to the end of the tine, what effect does this have on the wavelength and frequency of the sound the tine produces? Why?

Q16.14 An organist in a cathedral plays a loud chord and then releases the keys. The sound persists for a few seconds and gradually dies away. Why does it persist? What happens to the sound energy when the sound dies away?

Q16.15 Two loudspeakers, A and B, are driven by the same amplifier and emit sinusoidal waves in phase. The frequency of the waves emitted by each speaker is 860 Hz. Point P is 12.0 m from A and 13.4 m from B. Is the interference at P constructive or destructive? Give the reasoning behind your answer.

Q16.16 Two vibrating tuning forks have identical frequencies, but one is stationary and the other is mounted at the rim of a rotating platform. What does a listener hear? Explain.

Q16.17 A large church has part of the organ in the front of the church and part in the back. A person walking rapidly down the aisle while both segments are playing at once reports that the two segments sound out of tune. Why?

Q16.18 A sound source and a listener are both at rest on the earth, but a strong wind is blowing from the source toward the listener. Is there a Doppler effect? Why or why not?

Q16.19 Can you think of circumstances in which a Doppler effect would be observed for surface waves in water? For elastic waves propagating in a body of water deep below the surface? If so, describe the circumstances and explain your reasoning. If not, explain why not.

Q16.20 Stars other than our sun normally appear featureless when viewed through telescopes. Yet astronomers can readily use the light from these stars to determine that they are rotating and even measure the speed of their surface. How do you think they can do this?

Q16.21 If you wait at a railroad crossing as a train approaches and passes, you hear a Doppler shift in its sound. But if you listen closely, you hear that the change in frequency is continuous; it does not suddenly go from one high frequency to another low frequency. Instead the frequency *smoothly* (but rather quickly) changes from high to low as the train passes. Why does this smooth change occur?

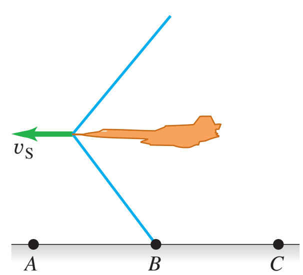
Q16.22 In case 1, a source of sound approaches a stationary observer at speed u. In case 2, the observer moves toward the stationary source at the same speed u. If the source is always producing the same frequency sound, will the observer hear the same frequency in both cases, since the relative speed is the same each time? Why or why not?

Q16.23 Does an aircraft make a sonic boom only at the instant its speed exceeds Mach 1? Explain.

Q16.24 If you are riding in a supersonic aircraft, what do you hear? Explain. In particular, do you hear a continuous sonic boom? Why or why not?

Q16.25 A jet airplane is flying at a constant altitude at a steady speed v_S greater than the speed of sound. Describe what observers at points *A*, *B*, and *C* hear at the instant shown in **Fig. Q16.25**, when the shock wave has just reached point *B*. Explain.

Figure **Q16.25**



EXERCISES

Unless indicated otherwise, assume the speed of sound in air to be $v = 344$ m/s.

Section 16.1 Sound Waves

16.1 • Example 16.1 (Section 16.1) showed that for sound waves in air with frequency 1000 Hz, a displacement amplitude of 1.2×10^{-8} m produces a pressure amplitude of 3.0×10^{-2} Pa. (a) What is the wavelength of these waves? (b) For 1000 Hz waves in air, what displacement amplitude would be needed for the pressure amplitude to be at the pain threshold, which is 30 Pa? (c) For what wavelength and frequency will waves with a displacement amplitude of 1.2×10^{-8} m produce a pressure amplitude of 1.5×10^{-3} Pa?

16.2 • A loud factory machine produces sound having a displacement amplitude of 1.00 μm, but the frequency of this sound can be adjusted. In order to prevent ear damage to the workers, the maximum pressure amplitude of the sound waves is limited to 10.0 Pa. Under the conditions of this factory, the bulk modulus of air is 1.42×10^5 Pa. What is the highest-frequency sound to which this machine can be adjusted without exceeding the prescribed limit? Is this frequency audible to the workers?

16.3 • Consider a sound wave in air that has displacement amplitude 0.0200 mm. Calculate the pressure amplitude for frequencies of (a) 150 Hz; (b) 1500 Hz; (c) 15,000 Hz. In each case compare the result to the pain threshold, which is 30 Pa.

16.4 • BIO **Ultrasound and Infrasound.** (a) **Whale communication.** Blue whales apparently communicate with each other using sound of frequency 17 Hz, which can be heard nearly 1000 km away in the ocean. What is the wavelength of such a sound in seawater, where the speed of sound is 1531 m/s? (b) **Dolphin clicks.** One type of sound that dolphins emit is a sharp click of wavelength 1.5 cm in the ocean. What is the frequency of such clicks? (c) **Dog whistles.** One brand of dog whistles claims a frequency of 25 kHz for its product. What is the wavelength of this sound? (d) **Bats.** While bats emit a wide variety of sounds, one type emits pulses of sound having a frequency between 39 kHz and 78 kHz. What is the range of wavelengths of this sound? (e) **Sonograms.** Ultrasound is used to view the interior of the body, much as x rays are utilized. For sharp imagery, the wavelength of the sound should be around one-fourth (or less) the size of the objects to be viewed. Approximately what frequency of sound is needed to produce a clear image of a tumor that is 1.0 mm across if the speed of sound in the tissue is 1550 m/s?

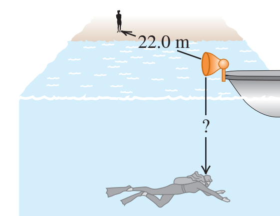
Section 16.2 Speed of Sound Waves

16.5 •• A 60.0-m-long brass rod is struck at one end. A person at the other end hears two sounds as a result of two longitudinal waves, one traveling in the metal rod and the other traveling in air. What is the time interval between the two sounds? (The speed of sound in air is 344 m/s; see Tables 11.1 and 12.1 for relevant information about brass.)

16.6 • (a) In a liquid with density 1300 kg/m^3, longitudinal waves with frequency 400 Hz are found to have wavelength 8.00 m. Calculate the bulk modulus of the liquid. (b) A metal bar with a length of 1.50 m has density 6400 kg/m^3. Longitudinal sound waves take 3.90×10^{-4} s to travel from one end of the bar to the other. What is Young's modulus for this metal?

16.7 • A submerged scuba diver hears the sound of a boat horn directly above her on the surface of the lake. At the same time, a friend on dry land 22.0 m from the boat also hears the horn (**Fig. E16.7**). The horn is 1.2 m above the surface of the water. What is the distance (labeled "?") from the horn to the diver? Both air and water are at 20°C.

Figure **E16.7**



16.8 • At a temperature of 27.0°C, what is the speed of longitudinal waves in (a) hydrogen (molar mass 2.02 g/mol); (b) helium (molar mass 4.00 g/mol); (c) argon (molar mass 39.9 g/mol)? See Table 19.1 for values of γ. (d) Compare your answers for parts (a), (b), and (c) with the speed in air at the same temperature.

16.9 • An oscillator vibrating at 1250 Hz produces a sound wave that travels through an ideal gas at 325 m/s when the gas temperature is 22.0°C. For a certain experiment, you need to have the same oscillator produce sound of wavelength 28.5 cm in this gas. What should the gas temperature be to achieve this wavelength?

16.10 •• CALC (a) Show that the fractional change in the speed of sound (dv/v) due to a very small temperature change dT is given by $dv/v = \frac{1}{2}dT/T$. [*Hint:* Start with Eq. (16.10).] (b) The speed of sound in air at 20°C is found to be 344 m/s. Use the result in part (a) to find the change in the speed of sound for a 1.0°C change in air temperature.

Section 16.3 Sound Intensity

16.11 •• BIO **Energy Delivered to the Ear.** Sound is detected when a sound wave causes the tympanic membrane (the eardrum) to vibrate. Typically, the diameter of this membrane is about 8.4 mm in humans. (a) How much energy is delivered to the eardrum each second when someone whispers (20 dB) a secret in your ear? (b) To comprehend how sensitive the ear is to very small amounts of energy, calculate how fast a typical 2.0 mg mosquito would have to fly (in mm/s) to have this amount of kinetic energy.

16.12 • (a) By what factor must the sound intensity be increased to raise the sound intensity level by 13.0 dB? (b) Explain why you don't need to know the original sound intensity.

16.13 •• **Eavesdropping!** You are trying to overhear a juicy conversation, but from your distance of 15.0 m, it sounds like only an average whisper of 20.0 dB. How close should you move to the chatterboxes for the sound level to be 60.0 dB?

16.14 • A small source of sound waves emits uniformly in all directions. The total power output of the source is *P*. By what factor must *P* increase if the sound intensity level at a distance of 20.0 m from the source is to increase 5.00 dB?

16.15 • A sound wave in air at 20°C has a frequency of 320 Hz and a displacement amplitude of 5.00×10^{-3} mm. For this sound wave calculate the (a) pressure amplitude (in Pa); (b) intensity (in W/m^2); (c) sound intensity level (in decibels).

16.16 •• You live on a busy street, but as a music lover, you want to reduce the traffic noise. (a) If you install special sound-reflecting windows that reduce the sound intensity level (in dB) by 30 dB, by what fraction have you lowered the sound intensity (in W/m^2)? (b) If, instead, you reduce the intensity by half, what change (in dB) do you make in the sound intensity level?

16.17 • BIO For a person with normal hearing, the faintest sound that can be heard at a frequency of 400 Hz has a pressure amplitude of about 6.0×10^{-5} Pa. Calculate the (a) intensity; (b) sound intensity level; (c) displacement amplitude of this sound wave at 20°C.

16.18 •• The intensity due to a number of independent sound sources is the sum of the individual intensities. (a) When four quadruplets cry simultaneously, how many decibels greater is the sound intensity level than when a single one cries? (b) To increase the sound intensity level again by the same number of decibels as in part (a), how many more crying babies are required?

16.19 • CP A baby's mouth is 30 cm from her father's ear and 1.50 m from her mother's ear. What is the difference between the sound intensity levels heard by the father and by the mother?

16.20 •• (a) If two sounds differ by 5.00 dB, find the ratio of the intensity of the louder sound to that of the softer one. (b) If one sound is 100 times as intense as another, by how much do they differ in sound intensity level (in decibels)? (c) If you increase the volume of your stereo so that the intensity doubles, by how much does the sound intensity level increase?

16.21 •• CP At point A, 3.0 m from a small source of sound that is emitting uniformly in all directions, the sound intensity level is 53 dB. (a) What is the intensity of the sound at A? (b) How far from the source must you go so that the intensity is one-fourth of what it was at A? (c) How far must you go so that the sound intensity level is one-fourth of what it was at A? (d) Does intensity obey the inverse-square law? What about sound intensity level?

16.22 •• The pattern of displacement nodes N and antinodes A in a pipe is *ANANANANANA* when the standing-wave frequency is 1710 Hz. The pipe contains air at 20°C (See Table 16.1.) (a) Is it an open or a closed (stopped) pipe? (b) Which harmonic is this? (c) What is the length of the pipe? (d) What is the fundamental frequency? (e) What would be the fundamental frequency of the pipe if it contained helium at 20°C?

Section 16.4 Standing Sound Waves and Normal Modes

16.23 • Standing sound waves are produced in a pipe that is 1.20 m long. For the fundamental and first two overtones, determine the locations along the pipe (measured from the left end) of the displacement nodes and the pressure nodes if (a) the pipe is open at both ends and (b) the pipe is closed at the left end and open at the right end.

16.24 • The fundamental frequency of a pipe that is open at both ends is 524 Hz. (a) How long is this pipe? If one end is now closed, find (b) the wavelength and (c) the frequency of the new fundamental.

16.25 • BIO The Human Voice. The human vocal tract is a pipe that extends about 17 cm from the lips to the vocal folds (also called "vocal cords") near the middle of your throat. The vocal folds behave rather like the reed of a clarinet, and the vocal tract acts like a stopped pipe. Estimate the first three standing-wave frequencies of the vocal tract. Use $v = 344$ m/s. (The answers are only an estimate, since the position of lips and tongue affects the motion of air in the vocal tract.)

16.26 •• BIO The Vocal Tract. Many professional singers have a range of $2\frac{1}{2}$ octaves or even greater. Suppose a soprano's range extends from A below middle C (frequency 220 Hz) up to E-flat above high C (frequency 1244 Hz). Although the vocal tract is complicated, we can model it as a resonating air column, like an organ pipe, that is open at the top and closed at the bottom. The column extends from the mouth down to the diaphragm in the chest cavity. Assume that the lowest note is the fundamental. How long is this column of air if $v = 354$ m/s? Does your result seem reasonable, on the basis of observations of your body?

16.27 • The longest pipe found in most medium-size pipe organs is 4.88 m (16 ft) long. What is the frequency of the note corresponding to the fundamental mode if the pipe is (a) open at both ends, (b) open at one end and closed at the other?

16.28 • Singing in the Shower. A pipe closed at both ends can have standing waves inside of it, but you normally don't hear them because little of the sound can get out. But you *can* hear them if you are *inside* the pipe, such as someone singing in the shower. (a) Show that the wavelengths of standing waves in a pipe of length L that is closed at both ends are $\lambda_n = 2L/n$ and the frequencies are given by $f_n = nv/2L = nf_1$, where $n = 1, 2, 3, \ldots$. (b) Modeling it as a pipe, find the frequency of the fundamental and the first two overtones for a shower 2.50 m tall. Are these frequencies audible?

16.29 •• The pattern of displacement nodes N and antinodes A in a pipe is *NANANANANA* when the standing-wave frequency is 1710 Hz. The pipe contains air at 20°C. (See Table 16.1.) (a) Is it an open or a closed (stopped) pipe? (b) Which harmonic is this? (c) What is the length of the pipe? (d) What is the fundamental frequency?

Section 16.5 Resonance and Sound

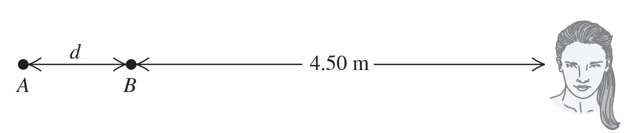
16.30 •• CP You have a stopped pipe of adjustable length close to a taut 62.0 cm, 7.25 g wire under a tension of 4110 N. You want to adjust the length of the pipe so that, when it produces sound at its fundamental frequency, this sound causes the wire to vibrate in its second *overtone* with very large amplitude. How long should the pipe be?

16.31 • You blow across the open mouth of an empty test tube and produce the fundamental standing wave in the 14.0-cm-long air column in the test tube, which acts as a stopped pipe. (a) What is the frequency of this standing wave? (b) What is the frequency of the fundamental standing wave in the air column if the test tube is half filled with water?

Section 16.6 Interference of Waves

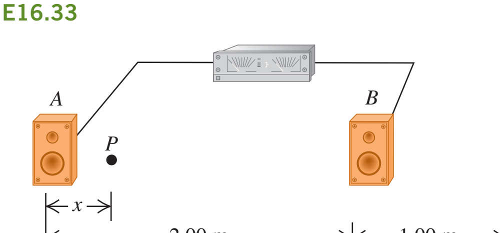
16.32 • Small speakers A and B are driven in phase at 725 Hz by the same audio oscillator. Both speakers start out 4.50 m from the listener, but speaker A is slowly moved away (**Fig. E16.32**). (a) At what distance d will the sound from the speakers first produce destructive interference at the listener's location? (b) If A is moved even farther away than in part (a), at what distance d will the speakers next produce destructive interference at the listener's location? (c) After A starts moving away from its original spot, at what distance d will the speakers first produce constructive interference at the listener's location?

Figure **E16.32**



16.33 • Two loudspeakers, A and B (**Fig. E16.33**), are driven by the same amplifier and emit sinusoidal waves in phase. Speaker B is 2.00 m to the right of speaker A. Consider point Q along the extension of the line connecting the speakers, 1.00 m to the right of speaker B. Both speakers emit sound waves that travel directly from the speaker to point Q. What is the lowest frequency for which (a) *constructive* interference occurs at point Q; (b) *destructive* interference occurs at point Q?

Figure **E16.33**



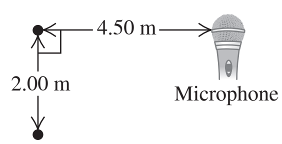
16.34 •• Two loudspeakers, A and B (see Fig. E16.33), are driven by the same amplifier and emit sinusoidal waves in phase. Speaker B is 2.00 m to the right of speaker A. The frequency of the sound waves produced by the loudspeakers is 206 Hz. Consider a point P between the speakers and along the line connecting them, a distance x to the right of A. Both speakers emit sound waves that travel directly from the speaker to point P. For what values of x will (a) *destructive* interference occur at P; (b) *constructive* interference occur at P? (c) Interference effects like those in parts (a) and (b) are almost never a factor in listening to home stereo equipment. Why not?

16.35 •• Two loudspeakers, A and B, are driven by the same amplifier and emit sinusoidal waves in phase. Speaker B is 12.0 m to the right of speaker A. The frequency of the waves emitted by each speaker is 688 Hz. You are standing between the speakers, along the line connecting them, and are at a point of constructive interference. How far must you walk toward speaker B to move to a point of destructive interference?

16.36 • Two loudspeakers, A and B, are driven by the same amplifier and emit sinusoidal waves in phase. The frequency of the waves emitted by each speaker is 172 Hz. You are 8.00 m from A. What is the closest you can be to B and be at a point of destructive interference?

16.37 •• Two small stereo speakers are driven in step by the same variable-frequency oscillator. Their sound is picked up by a microphone arranged as shown in **Fig. E16.37**. For what frequencies does their sound at the speakers produce (a) constructive interference and (b) destructive interference?

Figure **E16.37**



Section 16.7 Beats

16.38 •• Two guitarists attempt to play the same note of wavelength 64.8 cm at the same time, but one of the instruments is slightly out of tune and plays a note of wavelength 65.2 cm instead. What is the frequency of the beats these musicians hear when they play together?

16.39 •• **Tuning a Violin.** A violinist is tuning her instrument to concert A (440 Hz). She plays the note while listening to an electronically generated tone of exactly that frequency and hears a beat frequency of 3 Hz, which increases to 4 Hz when she tightens her violin string slightly. (a) What was the frequency of the note played by her violin when she heard the 3 Hz beats? (b) To get her violin perfectly tuned to concert A, should she tighten or loosen her string from what it was when she heard the 3 Hz beats?

16.40 •• Two organ pipes, open at one end but closed at the other, are each 1.14 m long. One is now lengthened by 2.00 cm. Find the beat frequency that they produce when playing together in their fundamentals.

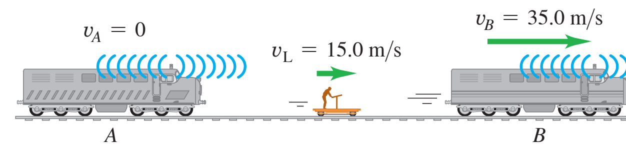
Section 16.8 The Doppler Effect

16.41 •• On the planet Arrakis a male ornithoid is flying toward his mate at 25.0 m/s while singing at a frequency of 1200 Hz. If the stationary female hears a tone of 1240 Hz, what is the speed of sound in the atmosphere of Arrakis?

16.42 • A railroad train is traveling at 25.0 m/s in still air. The frequency of the note emitted by the locomotive whistle is 400 Hz. What is the wavelength of the sound waves (a) in front of the locomotive and (b) behind the locomotive? What is the frequency of the sound heard by a stationary listener (c) in front of the locomotive and (d) behind the locomotive?

16.43 • Two train whistles, A and B, each have a frequency of 392 Hz. A is stationary and B is moving toward the right (away from A) at a speed of 35.0 m/s. A listener is between the two whistles and is moving toward the right with a speed of 15.0 m/s (**Fig. E16.43**). No wind is blowing. (a) What is the frequency from A as heard by the listener? (b) What is the frequency from B as heard by the listener? (c) What is the beat frequency detected by the listener?

Figure **E16.43**



16.44 • **Moving Source vs. Moving Listener.** (a) A sound source producing 1.00 kHz waves moves toward a stationary listener at one-half the speed of sound. What frequency will the listener hear? (b) Suppose instead that the source is stationary and the listener moves toward the source at one-half the speed of sound. What frequency does the listener hear? How does your answer compare to that in part (a)? Explain on physical grounds why the two answers differ.

16.45 • A swimming duck paddles the water with its feet once every 1.6 s, producing surface waves with this period. The duck is moving at constant speed in a pond where the speed of surface waves is 0.32 m/s, and the crests of the waves ahead of the duck are spaced 0.12 m apart. (a) What is the duck's speed? (b) How far apart are the crests behind the duck?

16.46 • A railroad train is traveling at 30.0 m/s in still air. The frequency of the note emitted by the train whistle is 352 Hz. What frequency is heard by a passenger on a train moving in the opposite direction to the first at 18.0 m/s and (a) approaching the first and (b) receding from the first?

16.47 • A car alarm is emitting sound waves of frequency 520 Hz. You are on a motorcycle, traveling directly away from the parked car. How fast must you be traveling if you detect a frequency of 490 Hz?

16.48 •• While sitting in your car by the side of a country road, you are approached by your friend, who happens to be in an identical car. You blow your car's horn, which has a frequency of 260 Hz. Your friend blows his car's horn, which is identical to yours, and you hear a beat frequency of 6.0 Hz. How fast is your friend approaching you?

16.49 • A police car is traveling due east at a speed of 15.0 m/s relative to the earth. You are in a convertible following behind the police car. Your car is also moving due east at 15.0 m/s relative to the earth, so the speed of the police car relative to you is zero. The siren of the police car is emitting sound of frequency 500 Hz. The speed of sound in the still air is 340 m/s. (a) What is the speed of the sound waves relative to you? (b) What is the wavelength of the sound waves at your location? (c) What frequency do you detect?

16.50 •• The siren of a fire engine that is driving northward at 30.0 m/s emits a sound of frequency 2000 Hz. A truck in front of this fire engine is moving northward at 20.0 m/s. (a) What is the frequency of the siren's sound that the fire engine's driver hears reflected from the back of the truck? (b) What wavelength would this driver measure for these reflected sound waves?

16.51 •• A stationary police car emits a sound of frequency 1200 Hz that bounces off a car on the highway and returns with a frequency of 1250 Hz. The police car is right next to the highway, so the moving car is traveling directly toward or away from it. (a) How fast was the moving car going? Was it moving toward or away from the police car? (b) What frequency would the police car have received if it had been traveling toward the other car at 20.0 m/s?

16.52 •• A stationary source emits sound waves of frequency f_s. There is no wind blowing. A device for detecting sound waves and measuring their observed frequency moves toward the source with speed v_L, and the observed frequency of the sound waves is f_L. The measurement is repeated for different values of v_L. You plot the results as f_L versus v_L and find that your data lie close to a straight line that has slope 1.75 m^{-1} and y-intercept 600.0 Hz. What are your experimental results for the speed of sound in the still air and for the frequency f_s of the source?

Section 16.9 Shock Waves

16.53 •• A jet plane flies overhead at Mach 1.70 and at a constant altitude of 1250 m. (a) What is the angle α of the shock-wave cone? (b) How much time after the plane passes directly overhead do you hear the sonic boom? Neglect the variation of the speed of sound with altitude.

16.54 • The shock-wave cone created by a space shuttle at one instant during its reentry into the atmosphere makes an angle of 58.0° with its direction of motion. The speed of sound at this altitude is 331 m/s. (a) What is the Mach number of the shuttle at this instant, and (b) how fast (in m/s and in mi/h) is it traveling relative to the atmosphere? (c) What would be its Mach number and the angle of its shock-wave cone if it flew at the same speed but at low altitude where the speed of sound is 344 m/s?

PROBLEMS

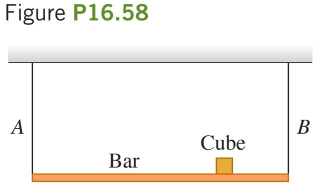
16.55 •• Use Eq. (16.10) and the information given in Example 16.4 to show that the speed of sound in air at 0°C is 332 m/s. (a) Using the first two terms of the power series expansion of $(1 + x)^n$ (see Appendix B), show that the speed of sound in air at Celsius temperature T_C is given approximately by $v = (332\text{ m/s})(1 + T_C/546)$. (b) Use the result in part (a) to calculate v at 20°C. Compare your result to the value given in Table 16.1. Is the expression in part (a) accurate at 20°C? (c) Do you expect the expression in part (a) to be accurate at 120°C? Explain. (*Hint:* Compare the second and third terms in the power series expansion.)

16.56 •• CP The sound from a trumpet radiates uniformly in all directions in 20°C air. At a distance of 5.00 m from the trumpet the sound intensity level is 52.0 dB. The frequency is 587 Hz. (a) What is the pressure amplitude at this distance? (b) What is the displacement amplitude? (c) At what distance is the sound intensity level 30.0 dB?

16.57 ••• CALC The air temperature over a lake decreases linearly with height after sunset, since air cools faster than water. (a) If the temperature at the surface is 25.00°C and the temperature at a height of 300.0 m is 5.000°C, how long does it take sound to rise 300.0 m directly upward? [*Hint:* Use Eq. (16.10) and integrate.] (b) At a height of 300.0 m, how far does sound travel horizontally in this same time interval? The change in wave speed with altitude due to the nocturnal temperature inversion over the lake gives rise to a change in direction of the sound waves. This phenomenon is called refraction and will be discussed in detail for electromagnetic waves in Chapter 33.

16.58 •• CP A uniform 165 N bar is supported horizontally by two identical wires A and B (**Fig. P16.58**). A small 185 N cube of lead is placed three-fourths of the way from A to B. The wires are each 75.0 cm long and have a mass of 5.50 g. If both of them are simultaneously plucked at the center, what is the frequency of the beats that they will produce when vibrating in their fundamental?

Figure **P16.58**

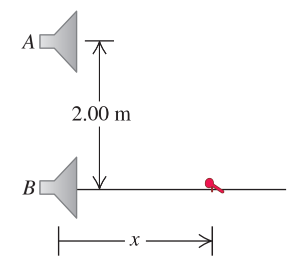


16.59 • An organ pipe has two successive harmonics with frequencies 1372 and 1764 Hz. (a) Is this an open or a stopped pipe? Explain. (b) What two harmonics are these? (c) What is the length of the pipe?

16.60 •• A Kundt's tube is filled with helium gas. The speed of sound for helium at 20°C is given in Table 16.1. A tuning fork that produces sound waves with frequency 1200 Hz is used to set up standing waves inside the tube. You measure the node-to-node distance to be 47.0 cm. What is the temperature of the helium gas in the tube?

16.61 •• Two identical loudspeakers are located at points A and B, 2.00 m apart. The loudspeakers are driven by the same amplifier and produce sound waves with a frequency of 784 Hz. Take the speed of sound in air to be 344 m/s. A small microphone is moved out from point B along a line perpendicular to the line connecting A and B (line BC in **Fig. P16.61**). (a) At what distances from B will there be *destructive* interference? (b) At what distances from B will there be *constructive* interference? (c) If the frequency is made low enough, there will be no positions along the line BC at which destructive interference occurs. How low must the frequency be for this to be the case?

Figure **P16.61**



16.62 •• A bat flies toward a wall, emitting a steady sound of frequency 1.70 kHz. This bat hears its own sound plus the sound reflected by the wall. How fast should the bat fly in order to hear a beat frequency of 8.00 Hz?

16.63 •• The sound source of a ship's sonar system operates at a frequency of 18.0 kHz. The speed of sound in water (assumed to be at a uniform 20°C) is 1482 m/s. (a) What is the wavelength of the waves emitted by the source? (b) What is the difference in frequency between the directly radiated waves and the waves reflected from a whale traveling directly toward the ship at 4.95 m/s? The ship is at rest in the water.

16.64 •• Consider a thunderstorm with a flash of lightning followed by a crash of thunder. (a) Estimate the time delay between the lightning flash and the sound of the thunder. (b) Determine the distance sound travels in that time, which provides an estimate of the distance to the lightning flash. (c) By comparing to the sounds listed in Table 16.2, estimate for your location the average intensity level of the sound in decibels and calculate the intensity in W/m^2. (d) Assume the sound was generated at the site of the lightning flash and assume the sound was transmitted uniformly in all directions. Use the estimated distance to estimate the average sound power generated by the thunder. (e) Estimate the duration of the thunderclap. Multiply your estimate by the average power to determine the sound energy released by a lightning strike.

16.65 ••• CP Suppose that you are at a bowling alley. (a) By comparing to the sounds listed in Table 16.2, for your location estimate the sound intensity level in decibels and the intensity in W/m^2 of the crashing pins after a well-executed strike. (b) Using 18 m as the length of a bowling alley, determine the average power associated with that sound. (c) Estimate the duration of the crashing sound. Multiply your estimate by the average power to obtain the sound energy released in the strike. (d) Estimate the time it took the bowling ball to travel down the alley prior to the strike. Use that time to estimate the speed of the ball. (e) Assume the ball has a mass of 6.4 kg. Account for both translational and rotational motions to estimate the kinetic energy of the bowling ball immediately prior to the strike. (f) What fraction of the ball's energy was converted to sound?

16.66 ••• BIO **Ultrasound in Medicine.** A 2.00 MHz sound wave travels through a pregnant woman's abdomen and is reflected from the fetal heart wall of her unborn baby. The heart wall is moving toward the sound receiver as the heart beats. The reflected sound is then mixed with the transmitted sound, and 72 beats per second are detected. The speed of sound in body tissue is 1500 m/s. Calculate the speed of the fetal heart wall at the instant this measurement is made.

16.67 ••• **BIO** Horseshoe bats (genus *Rhinolophus*) emit sounds from their nostrils and then listen to the frequency of the sound reflected from their prey to determine the prey's speed. (The "horseshoe" that gives the bat its name is a depression around the nostrils that acts like a focusing mirror, so that the bat emits sound in a narrow beam like a flashlight.) A *Rhinolophus* flying at speed v_{bat} emits sound of frequency f_{bat}; the sound it hears reflected from an insect flying toward it has a higher frequency f_{refl}. (a) Show that the speed of the insect is

$$v_{insect} = v\left[\frac{f_{refl}(v - v_{bat}) - f_{bat}(v + v_{bat})}{f_{refl}(v - v_{bat}) + f_{bat}(v + v_{bat})}\right]$$

where v is the speed of sound. (b) If $f_{bat} = 80.7$ kHz, $f_{refl} = 83.5$ kHz, and $v_{bat} = 3.9$ m/s, calculate the speed of the insect.

16.68 • **CP** A police siren of frequency f_{siren} is attached to a vibrating platform. The platform and siren oscillate up and down in simple harmonic motion with amplitude A_p and frequency f_p. (a) Find the maximum and minimum sound frequencies that you would hear at a position directly above the siren. (b) At what point in the motion of the platform is the maximum frequency heard? The minimum frequency? Explain.

16.69 •• **CP** A turntable 1.50 m in diameter rotates at 75 rpm. Two speakers, each giving off sound of wavelength 31.3 cm, are attached to the rim of the table at opposite ends of a diameter. A listener stands in front of the turntable. (a) What is the greatest beat frequency the listener will receive from this system? (b) Will the listener be able to distinguish individual beats?

16.70 •• **DATA** A long, closed cylindrical tank contains a diatomic gas that is maintained at a uniform temperature that can be varied. When you measure the speed of sound v in the gas as a function of the temperature T of the gas, you obtain these results:

T (°C)	−20.0	0.0	20.0	40.0	60.0	80.0
v (m/s)	324	337	349	361	372	383

(a) Explain how you can plot these results so that the graph will be well fit by a straight line. Construct this graph and verify that the plotted points do lie close to a straight line. (b) Because the gas is diatomic, $\gamma = 1.40$. Use the slope of the line in part (a) to calculate M, the molar mass of the gas. Express M in grams/mole. What type of gas is in the tank?

16.71 •• **DATA** A long tube contains air at a pressure of 1.00 atm and a temperature of 77.0°C. The tube is open at one end and closed at the other by a movable piston. A tuning fork that vibrates with a frequency of 500 Hz is placed near the open end. Resonance is produced when the piston is at distances 18.0 cm, 55.5 cm, and 93.0 cm from the open end. (a) From these values, what is the speed of sound in air at 77.0°C? (b) From the result of part (a), what is the value of γ? (c) These results show that a displacement antinode is slightly outside the open end of the tube. How far outside is it?

16.72 ••• **DATA** **Supernova!** (a) Equation (16.30) can be written as

$$f_R = f_S\left(1 - \frac{v}{c}\right)^{1/2}\left(1 + \frac{v}{c}\right)^{-1/2}$$

where c is the speed of light in vacuum, 3.00×10^8 m/s. Most objects move much slower than this (v/c is very small), so calculations made with Eq. (16.30) must be done carefully to avoid rounding errors. Use the binomial theorem to show that if $v \ll c$, Eq. (16.30) approximately reduces to $f_R = f_S[1 - (v/c)]$. (b) The gas cloud known as the Crab Nebula can be seen with even a small telescope. It is the remnant of a *supernova,* a cataclysmic explosion of a star. (The explosion was seen on the earth on July 4, 1054 C.E.) Its streamers glow with the characteristic red color of heated hydrogen gas. In a laboratory on the earth, heated hydrogen produces red light with frequency 4.568×10^{14} Hz; the red light received from streamers in the Crab Nebula that are pointed toward the earth has frequency 4.586×10^{14} Hz. Estimate the speed with which the outer edges of the Crab Nebula are expanding. Assume that the speed of the center of the nebula relative to the earth is negligible. (c) Assuming that the expansion speed of the Crab Nebula has been constant since the supernova that produced it, estimate the diameter of the Crab Nebula. Give your answer in meters and in light-years. (d) The angular diameter of the Crab Nebula as seen from the earth is about 5 arc-minutes $\left(1 \text{ arc-minute} = \frac{1}{60} \text{ degree}\right)$. Estimate the distance (in light-years) to the Crab Nebula, and estimate the year in which the supernova actually took place.

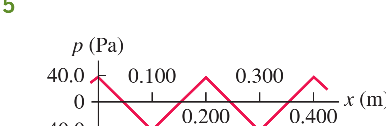
16.73 ••• **CP** A one-string Aeolian harp is constructed by attaching the lower end of a 1.56-m-long rigid and uniform pole with a mass of 8.00 kg to a tree at a pivot and then using a light rope to hang a 39.0-cm-long hollow steel tube with a mass of 4.00 kg from its upper end. A horizontal wire between the pole and the tree is attached to the pole at a height h above the pivot and holds the pole at a 45.0° angle with the vertical. The wire is uniform and has linear mass density $\mu = 1.40$ g/m. When the wind blows, the wire resonates at its fundamental frequency. The height h is chosen such that the sound emitted by the wire stimulates the fundamental standing wave in the tube, which harmonizes with the resonating wire. (a) What is the frequency of the note produced by this instrument? (b) At what height h should the wire be placed? (*Hint:* The pole is in static equilibrium, so the net torque on it is zero and this determines the tension in the wire as a function of h.)

16.74 ••• Two powerful speakers, separated by 15.00 m, stand on the floor in front of the stage in a large amphitheater. An aisle perpendicular to the stage is directly in front of one of the speakers and extends 50.00 m to an exit door at the back of the amphitheater. (a) If the speakers produce in-phase, coherent 440 Hz tones, at how many points along the aisle is the sound minimal? (b) What is the distance between the farthest such point and the door at the back of the aisle? (c) Suppose the coherent sound emitted from both speakers is a linear superposition of a 440 Hz tone and another tone with frequency f. What is the smallest value of f so that minimal sound is heard at any point where the 440 Hz sound is minimal? (d) At how many additional points in the aisle is the 440 Hz tone present but the second tone is minimal? (e) What is the distance from the closest of these points to the speaker at the front of the aisle?

CHALLENGE PROBLEMS

16.75 ••• **CALC** **Figure P16.75** shows the pressure fluctuation p of a nonsinusoidal sound wave as a function of x for $t = 0$. The wave is traveling in the $+x$-direction. (a) Graph the pressure fluctuation p as a function of t for $x = 0$. Show at least two cycles of oscillation. (b) Graph the displacement y in this sound wave as a function of x at $t = 0$. At $x = 0$, the displacement at $t = 0$ is zero. Show at least two wavelengths of the wave. (c) Graph the displacement y as a function of t for $x = 0$. Show at least two cycles of oscillation. (d) Calculate the maximum velocity and the maximum acceleration of an element of the air through which this sound wave is traveling. (e) Describe how the cone of a loudspeaker must move as a function of time to produce the sound wave in this problem.

Figure **P16.75**



16.76 ••• CP Longitudinal Waves on a Spring. A long spring such as a Slinky™ is often used to demonstrate longitudinal waves. (a) Show that if a spring that obeys Hooke's law has mass m, length L, and force constant k', the speed of longitudinal waves on the spring is $v = L\sqrt{k'/m}$ (see Section 16.2). (b) Evaluate v for a spring with $m = 0.250$ kg, $L = 2.00$ m, and $k' = 1.50$ N/m.

MCAT-STYLE PASSAGE PROBLEMS

BIO Ultrasound Imaging. A typical ultrasound transducer used for medical diagnosis produces a beam of ultrasound with a frequency of 1.0 MHz. The beam travels from the transducer through tissue and partially reflects when it encounters different structures in the tissue. The same transducer that produces the ultrasound also detects the reflections. The transducer emits a short pulse of ultrasound and waits to receive the reflected echoes before emitting the next pulse. By measuring the time between the initial pulse and the arrival of the reflected signal, we can use the speed of ultrasound in tissue, 1540 m/s, to determine the distance from the transducer to the structure that produced the reflection.

As the ultrasound beam passes through tissue, the beam is attenuated through absorption. Thus deeper structures return weaker echoes. A typical attenuation in tissue is -100 dB/m · MHz; in bone it is -500 dB/m · MHz. In determining attenuation, we take the reference intensity to be the intensity produced by the transducer.

16.77 If the deepest structure you wish to image is 10.0 cm from the transducer, what is the maximum number of pulses per second that can be emitted? (a) 3850; (b) 7700; (c) 15,400; (d) 1,000,000.

16.78 After a beam passes through 10 cm of tissue, what is the beam's intensity as a fraction of its initial intensity from the transducer? (a) 1×10^{-11}; (b) 0.001; (c) 0.01; (d) 0.1.

16.79 Because the speed of ultrasound in bone is about twice the speed in soft tissue, the distance to a structure that lies beyond a bone can be measured incorrectly. If a beam passes through 4 cm of tissue, then 2 cm of bone, and then another 1 cm of tissue before echoing off a cyst and returning to the transducer, what is the difference between the true distance to the cyst and the distance that is measured by assuming the speed is always 1540 m/s? Compared with the measured distance, the structure is actually (a) 1 cm farther; (b) 2 cm farther; (c) 1 cm closer; (d) 2 cm closer.

16.80 In some applications of ultrasound, such as its use on cranial tissues, large reflections from the surrounding bones can produce standing waves. This is of concern because the large pressure amplitude in an antinode can damage tissues. For a frequency of 1.0 MHz, what is the distance between antinodes in tissue? (a) 0.38 mm; (b) 0.75 mm; (c) 1.5 mm; (d) 3.0 mm.

16.81 For cranial ultrasound, why is it advantageous to use frequencies in the kHZ range rather than the MHz range? (a) The antinodes of the standing waves will be closer together at the lower frequencies than at the higher frequencies; (b) there will be no standing waves at the lower frequencies; (c) cranial bones will attenuate the ultrasound more at the lower frequencies than at the higher frequencies; (d) cranial bones will attenuate the ultrasound less at the lower frequencies than at the higher frequencies.

ANSWERS

Chapter Opening Question ?

(iv) Equation (16.10) in Section 16.2 says that the speed of sound in a gas depends on the temperature and on the kind of gas (through the ratio of heat capacities and the molar mass). Winter air in the mountains has a lower temperature than summer air at sea level, but they have essentially the same composition. Hence the lower temperature alone explains the slower speed of sound in winter in the mountains.

Key Example √ARIATION Problems

VP16.9.1 (a) $p_{max} = 6.74 \times 10^{-3}$ Pa (b) p_{max} is unchanged
VP16.9.2 (a) 3.16×10^{-4} W/m^2 (b) 63.1
VP16.9.3 (a) 102 dB (b) 90 dB
VP16.9.4 (a) 5.02×10^{-2} Pa (b) 86.0 dB
VP16.12.1 (a) 0.782 m (b) 0.130 m
VP16.12.2 (a) $n = 1$ (the fundamental frequency) (b) $n = 3$ (the third harmonic) (c) none
VP16.12.3 (a) 224 m/s (b) 0.521 m
VP16.12.4 (a) $f = 749$ Hz, $\lambda = 1.33$ m (b) $f = 749$ Hz, $\lambda = 0.459$ m (c) 1.15 m
VP16.18.1 (a) $f = 3.03 \times 10^3$ Hz, $\lambda = 0.112$ m (b) $f = 2.60 \times 10^3$ Hz, $\lambda = 0.131$ m
VP16.18.2 (a) $\lambda = 0.773$ m (b) $f = 453$ Hz, $\lambda = 0.773$ m (c) $f = 427$ Hz, $\lambda = 0.773$ m
VP16.18.3 (a) 1.22×10^3 Hz (b) 1.50×10^3 Hz
VP16.18.4 (a) 481 Hz (b) 29.9 m/s

Bridging Problem

(a) $180° = \pi$ rad
(b) A alone: $I = 3.98 \times 10^{-6}$ W/m^2, $\beta = 66.0$ dB;
B alone: $I = 5.31 \times 10^{-7}$ W/m^2, $\beta = 57.2$ dB
(c) $I = 1.60 \times 10^{-6}$ W/m^2, $\beta = 62.1$ dB

? At a steelworks, molten iron is heated to 1500° Celsius to remove impurities. It is most accurate to say that the molten iron contains a large amount of (i) temperature; (ii) heat; (iii) energy; (iv) two of these; (v) all three of these.

17 Temperature and Heat

Whether it's a sweltering summer day or a frozen midwinter night, your body needs to be kept at a nearly constant temperature. It has effective temperature-control mechanisms, but sometimes it needs help. On a hot day you wear less clothing to improve heat transfer from your body to the air and for better cooling by evaporation of perspiration. On a cold day you may sit by a roaring fire to absorb the energy that it radiates. The concepts in this chapter will help you understand the basic physics of keeping warm or cool.

The terms "temperature" and "heat" are often used interchangeably in everyday language. In physics, however, these two terms have very different meanings. In this chapter we'll define temperature in terms of how it's measured and see how temperature changes affect the dimensions of objects. We'll see that heat refers to energy transfer caused by temperature differences only and learn how to calculate and control such energy transfers.

Our emphasis in this chapter is on the concepts of temperature and heat as they relate to *macroscopic* objects such as cylinders of gas, ice cubes, and the human body. In Chapter 18 we'll look at these same concepts from a *microscopic* viewpoint in terms of the behavior of individual atoms and molecules. These two chapters lay the groundwork for the subject of **thermodynamics,** the study of energy transformations involving heat, mechanical work, and other aspects of energy and how these transformations relate to the properties of matter. Thermodynamics forms an indispensable part of the foundation of physics, chemistry, and the life sciences, and its applications turn up in such places as car engines, refrigerators, biochemical processes, and the structure of stars. We'll explore the key ideas of thermodynamics in Chapters 19 and 20.

LEARNING OUTCOMES

In this chapter, you'll learn...

- **17.1** The meaning of thermal equilibrium, and what thermometers really measure.
- **17.2** How different types of thermometers function.
- **17.3** The physics behind the absolute, or Kelvin, temperature scale.
- **17.4** How the dimensions of an object change as a result of a temperature change.
- **17.5** The meaning of heat, and how it differs from temperature.
- **17.6** How to do calculations that involve heat flow, temperature changes, and changes of phase.
- **17.7** How heat is transferred by conduction, convection, and radiation.

You'll need to review...

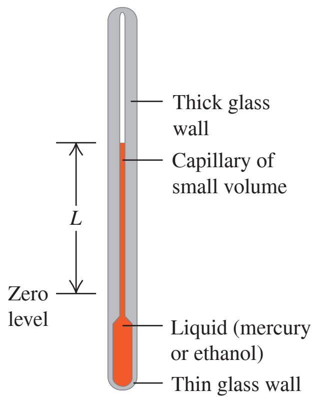
- **11.4** Stress and strain.
- **12.2** Measuring pressure.
- **14.4** Spring forces and interatomic forces.

17.1 TEMPERATURE AND THERMAL EQUILIBRIUM

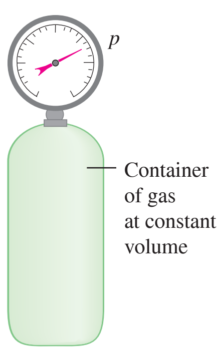
The concept of **temperature** is rooted in qualitative ideas based on our sense of touch. An object that feels "hot" usually has a higher temperature than a similar object that feels "cold." That's pretty vague, and the senses can be deceived. But many properties of matter that we can *measure*—including the length of a metal rod, steam pressure in a boiler, the ability of a wire to conduct an electric current, and the color of a very hot glowing object—depend on temperature.

Figure 17.1 Two devices for measuring temperature.

(a) Changes in temperature cause the liquid's volume to change.



(b) Changes in temperature cause the pressure of the gas to change.



Temperature is also related to the kinetic energies of the molecules of a material. In general this relationship is fairly complex, so it's not a good place to start in *defining* temperature. In Chapter 18 we'll look at the relationship between temperature and the energy of molecular motion for an ideal gas. However, we can define temperature and heat independently of any detailed molecular picture. In this section we'll develop a *macroscopic* definition of temperature.

To use temperature as a measure of hotness or coldness, we need to construct a temperature scale. To do this, we can use any measurable property of a system that varies with its "hotness" or "coldness." **Figure 17.1a** shows a familiar system that is used to measure temperature. When the system becomes hotter, the colored liquid (usually mercury or ethanol) expands and rises in the tube, and the value of L increases. Another simple system is a quantity of gas in a constant-volume container (Fig. 17.1b). The pressure p, measured by the gauge, increases or decreases as the gas becomes hotter or colder. A third example is the electrical resistance R of a conducting wire, which also varies when the wire becomes hotter or colder. Each of these properties gives us a number (L, p, or R) that varies with hotness and coldness, so each property can be used to make a **thermometer.**

To measure the temperature of an object, you place the thermometer in contact with the object. If you want to know the temperature of a cup of hot coffee, you stick the thermometer in the coffee; as the two interact, the thermometer becomes hotter and the coffee cools off a little. After the thermometer settles down to a steady value, you read the temperature. The system has reached an *equilibrium* condition, in which the interaction between the thermometer and the coffee causes no further change in the system. We call this a state of **thermal equilibrium.**

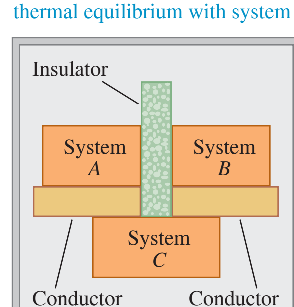
If two systems are separated by an insulating material or **insulator** such as wood, plastic foam, or fiberglass, they influence each other more slowly. Camping coolers are made with insulating materials to delay the cold food inside from warming up and attaining thermal equilibrium with the hot summer air outside. An *ideal insulator* is an idealized material that permits no interaction at all between the two systems. It prevents the systems from attaining thermal equilibrium if they aren't in thermal equilibrium at the start. Real insulators, like those in camping coolers, aren't ideal, so the contents of the cooler will warm up eventually. But an ideal insulator is nonetheless a useful idealization, like a massless rope or a frictionless incline.

The Zeroth Law of Thermodynamics

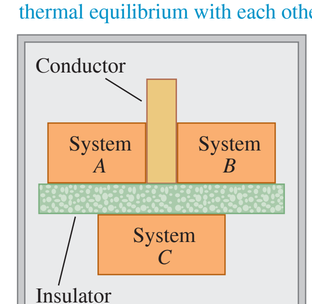
We can discover an important property of thermal equilibrium by considering three systems, A, B, and C, that initially are not in thermal equilibrium (**Fig. 17.2**). We surround them with an ideal insulating box so that they cannot interact with anything except each other. We separate systems A and B with an ideal insulating wall (the green slab in Fig. 17.2a), but we let system C interact with both systems A and B. We show this interaction in the figure by a yellow slab representing a thermal **conductor,** a material that *permits* thermal interactions through it. We wait until thermal equilibrium is attained; then A and B are each in thermal equilibrium with C. But are they in thermal equilibrium *with each other?*

Figure 17.2 The zeroth law of thermodynamics.

(a) If systems A and B are each in thermal equilibrium with system C ...



(b) ... then systems A and B are in thermal equilibrium with each other.



To find out, we separate system C from systems A and B with an ideal insulating wall (Fig. 17.2b), then replace the insulating wall between A and B with a *conducting* wall that lets A and B interact. What happens? Experiment shows that *nothing* happens; there are no additional changes to A or B. We can summarize this result as follows:

ZEROTH LAW OF THERMODYNAMICS If C is initially in thermal equilibrium with both A and B, then A and B are also in thermal equilibrium with each other.

(The importance of the zeroth law was recognized only after the first, second, and third laws of thermodynamics had been named. Since it is fundamental to all of them, the name "zeroth" seemed appropriate. We'll learn about the other laws of thermodynamics in Chapters 19 and 20.)

Now suppose system C is a thermometer, such as the liquid-in-tube system of Fig. 17.1a. In Fig. 17.2a the thermometer C is in contact with both A and B. In thermal equilibrium, when the thermometer reading reaches a stable value, the thermometer measures the temperature of both A and B; hence both A and B have the *same* temperature. Experiment shows that thermal equilibrium isn't affected by adding or removing insulators, so the reading of thermometer C wouldn't change if it were in contact only with A or only with B. We conclude:

CONDITION FOR THERMAL EQUILIBRIUM Two systems are in thermal equilibrium if and only if they have the same temperature.

This is what makes a thermometer useful; a thermometer actually measures *its own* temperature, but when a thermometer is in thermal equilibrium with another object, the temperatures must be equal. When the temperatures of two systems are different, they *cannot* be in thermal equilibrium.

TEST YOUR UNDERSTANDING OF SECTION 17.1 You put a thermometer in a pot of hot water and record the reading. What temperature have you recorded? (i) The temperature of the water; (ii) the temperature of the thermometer; (iii) an equal average of the temperatures of the water and thermometer; (iv) a weighted average of the temperatures of the water and thermometer, with more emphasis on the temperature of the water; (v) a weighted average of the water and thermometer, with more emphasis on the temperature of the thermometer.

ANSWER

(ii) A liquid-in-tube thermometer actually measures its own temperature. If the thermometer stays in the hot water long enough, it will come to thermal equilibrium with the water and its temperature will be the same as that of the water.

17.2 THERMOMETERS AND TEMPERATURE SCALES

To make the liquid-in-tube device shown in Fig. 17.1a into a useful thermometer, we need to mark a scale on the tube wall with numbers on it. Suppose we label the thermometer's liquid level at the freezing temperature of pure water "zero" and the level at the boiling temperature "100," and divide the distance between these two points into 100 equal intervals called *degrees*. The result is the **Celsius temperature scale** (formerly called the *centigrade* scale in English-speaking countries). The Celsius temperature for a state colder than freezing water is a negative number. The Celsius scale is used, both in everyday life and in science and industry, almost everywhere in the world.

Another common type of thermometer uses a *bimetallic strip,* made by bonding strips of two different metals together (**Fig. 17.3a**). When the temperature of the composite strip increases, one metal expands more than the other and the strip bends (Fig. 17.3b). This strip is usually formed into a spiral, with the outer end anchored to the thermometer case and the inner end attached to a pointer (Fig. 17.3c). The pointer rotates in response to temperature changes.

Figure **17.3** Use of a bimetallic strip as a thermometer.

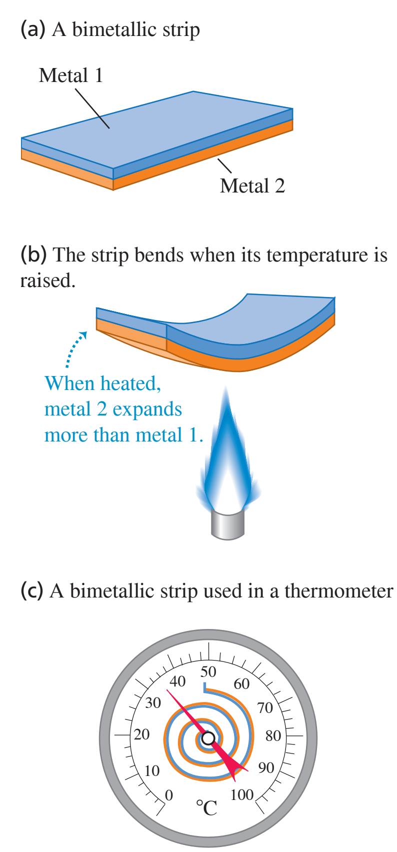


Figure 17.4 A temporal artery thermometer measures infrared radiation from the skin that overlies one of the important arteries in the head. Although the thermometer cover touches the skin, the infrared detector inside the cover does not.

In a *resistance thermometer* the changing electrical resistance of a coil of fine wire, a carbon cylinder, or a germanium crystal is measured. Resistance thermometers are usually more precise than most other types.

Some thermometers detect the amount of infrared radiation emitted by an object. (We'll see in Section 17.7 that *all* objects emit electromagnetic radiation, including infrared, as a consequence of their temperature.) One example is a *temporal artery thermometer* (**Fig. 17.4**). A nurse runs this over a patient's forehead in the vicinity of the temporal artery, and an infrared sensor in the thermometer measures the radiation from the skin. This device gives more accurate values of body temperature than do oral or ear thermometers.

In the **Fahrenheit temperature scale,** still used in the United States, the freezing temperature of water is 32°F and the boiling temperature is 212°F, both at standard atmospheric pressure. There are 180 degrees between freezing and boiling, compared to 100 on the Celsius scale, so one Fahrenheit degree represents only $\frac{100}{180}$, or $\frac{5}{9}$, as great a temperature change as one Celsius degree.

To convert temperatures from Celsius to Fahrenheit, note that a Celsius temperature T_C is the number of Celsius degrees above freezing; the number of Fahrenheit degrees above freezing is $\frac{9}{5}$ of this. But freezing on the Fahrenheit scale is at 32°F, so to obtain the actual Fahrenheit temperature T_F, multiply the Celsius value by $\frac{9}{5}$ and then add 32°. Symbolically,

$$\underbrace{T_F}_{\text{Fahrenheit temperature}} = \tfrac{9}{5}\underbrace{T_C}_{\text{Celsius temperature}} + 32^\circ \qquad (17.1)$$

To convert Fahrenheit to Celsius, solve this equation for T_C:

$$\underbrace{T_C}_{\text{Celsius temperature}} = \tfrac{5}{9}\left(\underbrace{T_F}_{\text{Fahrenheit temperature}} - 32^\circ\right) \qquad (17.2)$$

In words, subtract 32° to get the number of Fahrenheit degrees above freezing, and then multiply by $\frac{5}{9}$ to obtain the number of Celsius degrees above freezing—that is, the Celsius temperature.

We don't recommend memorizing Eqs. (17.1) and (17.2). Instead, understand the reasoning that led to them so that you can derive them on the spot when you need them, checking your reasoning with the relationship 100°C = 212°F.

It is useful to distinguish between an actual temperature and a temperature *interval* (a difference or change in temperature). An actual temperature of 20° is stated as 20°C (twenty degrees Celsius), and a temperature *interval* of 15° is 15C° (fifteen Celsius degrees). A beaker of water heated from 20°C to 35°C undergoes a temperature change of 15 C°.

CAUTION **Converting temperature differences** Keep in mind that Eqs. (17.1) and (17.2) apply to *temperatures,* not *temperature differences.* To convert a temperature difference in Fahrenheit degrees (F°) to one in Celsius degrees (C°), simply multiply by $\frac{5}{9}$; to convert a temperature difference in C° to one in F°, multiply by $\frac{9}{5}$.

TEST YOUR UNDERSTANDING OF SECTION 17.2 Which of the following types of thermometers have to be in thermal equilibrium with the object being measured in order to give accurate readings? (i) A bimetallic strip; (ii) a resistance thermometer; (iii) a temporal artery thermometer; (iv) both (i) and (ii); (v) all of (i), (ii), and (iii).

ANSWER

(iv) Both a bimetallic strip and a resistance thermometer measure their own temperature. For this to be equal to the temperature of the object being measured, the thermometer and object must be in thermal equilibrium. A temporal artery thermometer detects the infrared radiation from a person's skin; the detector and skin need not be at the same temperature.

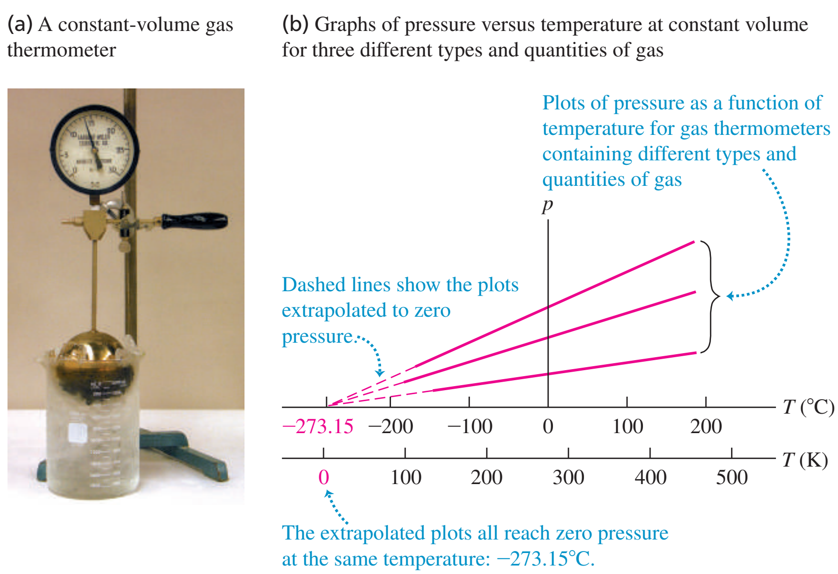


Figure 17.5 (a) Using a constant-volume gas thermometer to measure temperature. (b) The greater the amount of gas in the thermometer, the higher the graph of pressure p versus temperature T.

17.3 GAS THERMOMETERS AND THE KELVIN SCALE

When we calibrate two thermometers, such as a liquid-in-tube system and a resistance thermometer, so that they agree at 0°C and 100°C, they may not agree exactly at intermediate temperatures. Any temperature scale defined in this way always depends somewhat on the specific properties of the material used. Ideally, we would like to define a temperature scale that *doesn't* depend on the properties of a particular material. To establish a truly material-independent scale, we first need to develop some principles of thermodynamics. We'll return to this fundamental problem in Chapter 20. Here we'll discuss a thermometer that comes close to the ideal, the *constant-volume gas thermometer.*

The principle of a constant-volume gas thermometer is that the pressure of a gas at constant volume increases with temperature. We place a quantity of gas in a constant-volume container (**Fig. 17.5a**) and measure its pressure by one of the devices described in Section 12.2. To calibrate this thermometer, we measure the pressure at two temperatures, say 0°C and 100°C, plot these points on a graph, and draw a straight line between them. Then we can read from the graph the temperature corresponding to any other pressure. Figure 17.5b shows the results of three such experiments, each using a different type and quantity of gas.

By extrapolating this graph, we see that there is a hypothetical temperature, −273.15°C, at which the absolute pressure of the gas would become zero. This temperature turns out to be the *same* for many different gases (at least in the limit of very low gas density). We can't actually observe this zero-pressure condition. Gases liquefy and solidify at very low temperatures, and the proportionality of pressure to temperature no longer holds.

We use this extrapolated zero-pressure temperature as the basis for a temperature scale with its zero at this temperature. This is the **Kelvin temperature scale,** named for the British physicist Lord Kelvin (1824–1907). The units are the same size as those on the Celsius scale, but the zero is shifted so that $0\,\text{K} = -273.15°\text{C}$ and $273.15\,\text{K} = 0°\text{C}$ (Fig. 17.5b); that is,

$$\underset{\text{Kelvin temperature}}{T_K} = \underset{\text{Celsius temperature}}{T_C} + 273.15 \tag{17.3}$$

A common room temperature, 20°C (= 68°F), is 20 + 273.15, or about 293 K.

CAUTION Never say "degrees kelvin" In SI nomenclature, the temperature mentioned above is read "293 kelvins," not "degrees kelvin" (**Fig. 17.6**). We capitalize Kelvin when it refers to the temperature scale; however, the *unit* of temperature is the *kelvin,* which is not capitalized (but is nonetheless abbreviated as a capital K).

BIO APPLICATION Mammalian Body Temperatures Most mammals maintain body temperatures in the range from 36°C to 40°C (309 K to 313 K). A high metabolic rate warms the animal from within, and insulation (such as fur and body fat) slows heat loss.

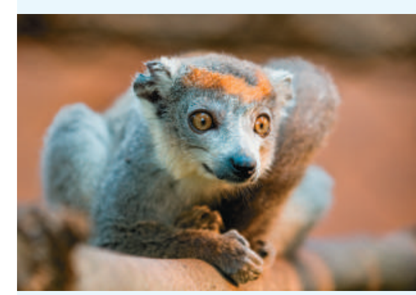
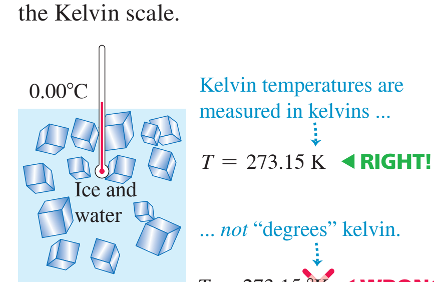


Figure 17.6 Correct and incorrect uses of the Kelvin scale.



EXAMPLE 17.1 Body temperature

You place a small piece of ice in your mouth. Eventually, the water all converts from ice at $T_1 = 32.00°\text{F}$ to body temperature, $T_2 = 98.60°\text{F}$. Express these temperatures in both Celsius degrees and kelvins, and find $\Delta T = T_2 - T_1$ in both cases.

IDENTIFY and SET UP Our target variables are stated above. We convert Fahrenheit temperatures to Celsius by using Eq. (17.2), and Celsius temperatures to Kelvin by using Eq. (17.3).

EXECUTE From Eq. (17.2), $T_1 = 0.00°\text{C}$ and $T_2 = 37.00°\text{C}$; then $\Delta T = T_2 - T_1 = 37.00\ \text{C}°$. To get the Kelvin temperatures, just add 273.15 to each Celsius temperature: $T_1 = 273.15$ K and $T_2 = 310.15$ K. The temperature difference is $\Delta T = T_2 - T_1 = 37.00$ K.

EVALUATE The Celsius and Kelvin scales have different zero points but the same size degrees. Therefore *any* temperature difference ΔT is the *same* on the Celsius and Kelvin scales. However, ΔT is *not* the same on the Fahrenheit scale; here, for example, $\Delta T = 66.60\ \text{F}°$.

KEYCONCEPT The Celsius and Kelvin scales have different zero points, but differences in temperature are the same in both scales: Increasing the temperature by 37.00 C° is the same as increasing it by 37.00 K.

The Kelvin Scale and Absolute Temperature

CAUTION Updating the Kelvin scale The definition of the Kelvin scale given here was accurate as of 2018. As of 2019, there is a new definition of this scale based on the definition of the joule and the value of the Boltzmann constant (which we'll introduce in Chapter 18). The change in definition has no effect, however, on the calculations in this textbook and calculations that you'll make.

The Celsius scale has two fixed points: the normal freezing and boiling temperatures of water. But we can define the Kelvin scale by using a gas thermometer with only a single reference temperature. Figure 17.5b shows that the pressure p in a gas thermometer is directly proportional to the Kelvin temperature. So we can define the ratio of any two Kelvin temperatures T_1 and T_2 as the ratio of the corresponding gas-thermometer pressures p_1 and p_2:

Definition of Kelvin scale: Ratio of two **temperatures in kelvins** ...

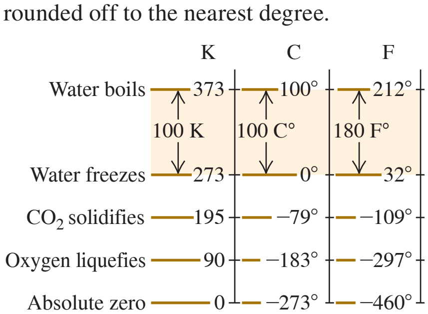
$$\frac{T_2}{T_1} = \frac{p_2}{p_1} \tag{17.4}$$

... equals ratio of corresponding **pressures in constant-volume gas thermometer**.

To complete the definition of T, we need only specify the Kelvin temperature of a single state. For reasons of precision and reproducibility, the state chosen is the *triple point* of water, the unique combination of temperature and pressure at which solid water (ice), liquid water, and water vapor can all coexist. It occurs at a temperature of 0.01°C and a water-vapor pressure of 610 Pa (about 0.006 atm). (This is the pressure of the *water,* not the gas pressure in the *thermometer*.) The triple-point temperature of water is *defined* to have the value $T_{\text{triple}} = 273.16$ K, corresponding to 0.01°C. From Eq. (17.4), if p_{triple} is the pressure in a gas thermometer at temperature T_{triple} and p is the pressure at some other temperature T, then T is given on the Kelvin scale by

$$T = T_{\text{triple}}\frac{p}{p_{\text{triple}}} = (273.16\ \text{K})\frac{p}{p_{\text{triple}}} \tag{17.5}$$

Figure 17.7 Relationships among Kelvin (K), Celsius (C), and Fahrenheit (F) temperature scales. Temperatures have been rounded off to the nearest degree.



Gas thermometers are impractical for everyday use. They are bulky and very slow to come to thermal equilibrium. They are used principally to establish high-precision standards and to calibrate other thermometers.

Figure 17.7 shows the relationships among the three temperature scales we have discussed. The Kelvin scale is called an **absolute temperature scale,** and its zero point [$T = 0\ \text{K} = -273.15°\text{C}$, the temperature at which $p = 0$ in Eq. (17.5)] is called **absolute zero.** At absolute zero a system of molecules (such as a quantity of a gas, a liquid, or a solid) has its *minimum* possible total energy (kinetic plus potential); because of quantum effects, it is *not* correct to say that all molecular motion ceases at absolute zero. In Chapter 20 we'll define more completely what we mean by absolute zero through thermodynamic principles that we'll develop in the next several chapters.

TEST YOUR UNDERSTANDING OF SECTION 17.3 Rank the following temperatures from highest to lowest: (i) 0.00°C; (ii) 0.00°F; (iii) 260.00 K; (iv) 77.00 K; (v) −180.00°C.

ANSWER

(i), (iii), (ii), (v), (iv) To compare these temperatures, convert them all to the Kelvin scale. For (i), the Kelvin temperature is $T_K = T_C + 273.15 = 0.00 + 273.15 = 273.15$ K; for (ii), $T_C = \frac{5}{9}(T_F - 32°) = \frac{5}{9}(0.00° - 32°) = -17.78°\text{C}$ and $T_K = T_C + 273.15 = -17.78 + 273.15 = 255.37$ K; for (iii), $T_K = 260.00$ K; for (iv), $T_K = 77.00$ K; and for (v), $T_K = T_C + 273.15 = -180.00 + 273.15 = 93.15$ K.

17.4 THERMAL EXPANSION

Most materials expand when their temperatures increase. Rising temperatures make the liquid expand in a liquid-in-tube thermometer (Fig. 17.1a) and bend bimetallic strips (Fig. 17.3b). A completely filled and tightly capped bottle of water cracks when it is heated, but you can loosen a metal jar lid by running hot water over it. These are all examples of *thermal expansion*.

Linear Expansion

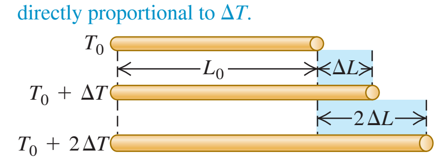
Suppose a solid rod has a length L_0 at some initial temperature T_0. When the temperature changes by ΔT, the length changes by ΔL. Experiments show that if ΔT is not too large (say, less than 100 C° or so), ΔL is *directly proportional* to ΔT (**Fig. 17.8a**). If two rods made of the same material have the same temperature change, but one is twice as long as the other, then the *change* in its length is also twice as great. Therefore ΔL must also be proportional to L_0 (Fig. 17.8b). We may express these relationships in an equation:

Linear thermal expansion:

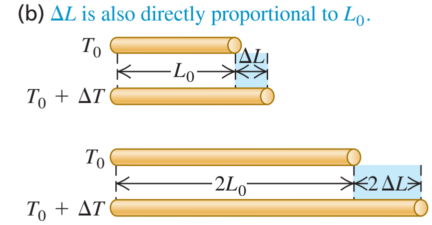
$$\underbrace{\Delta L}_{\text{Change in length}} = \underbrace{\alpha}_{\text{Coefficient of linear expansion}} \overbrace{L_0}^{\text{Original length}} \underbrace{\Delta T}_{\text{Temperature change}} \tag{17.6}$$

Figure 17.8 How the length of a rod changes with a change in temperature. (Length changes are exaggerated for clarity.)

(a) For moderate temperature changes, ΔL is directly proportional to ΔT.



(b) ΔL is also directly proportional to L_0.



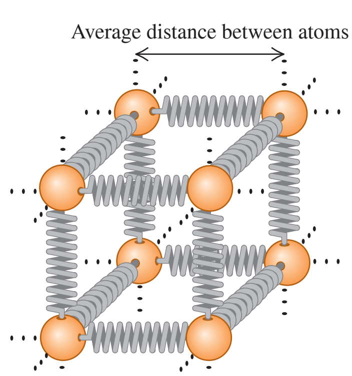
The constant α, which has different values for different materials, is called the **coefficient of linear expansion.** The units of α are K^{-1} or $(C°)^{-1}$. (Remember that a temperature interval is the same on the Kelvin and Celsius scales.) If an object has length L_0 at temperature T_0, then its length L at a temperature $T = T_0 + \Delta T$ is

$$L = L_0 + \Delta L = L_0 + \alpha L_0 \Delta T = L_0(1 + \alpha \Delta T) \tag{17.7}$$

For many materials, every linear dimension changes according to Eq. (17.6) or (17.7). Thus L could be the thickness of a rod, the side length of a square sheet, or the diameter of a hole. Some materials, such as wood or single crystals, expand differently in different directions. We won't consider this complication.

We can understand thermal expansion qualitatively on a molecular basis. Picture the interatomic forces in a solid as springs, as in **Fig. 17.9a**. (We explored the analogy between spring forces and interatomic forces in Section 14.4.) Each atom vibrates about its equilibrium position. When the temperature increases, the energy and amplitude of the vibration also increase. The interatomic spring forces are not symmetrical about the equilibrium position; they usually behave like a spring that is easier to stretch than to compress. As a result, when the amplitude of vibration increases, the *average* distance between atoms also increases (Fig. 17.9b). As the atoms get farther apart, every dimension increases.

(a) A model of the forces between neighboring atoms in a solid



(b) A graph of the "spring" potential energy $U(x)$

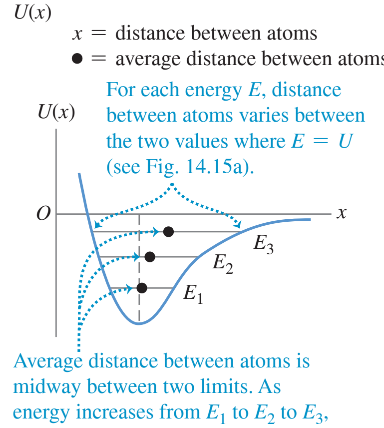


Figure 17.9 (a) We can model atoms in a solid as being held together by "springs" that are easier to stretch than to compress. (b) A graph of the "spring" potential energy $U(x)$ versus distance x between neighboring atoms is *not* symmetrical (compare Fig. 14.20b). As the energy increases and the atoms oscillate with greater amplitude, the average distance increases.

Figure 17.10 When an object undergoes thermal expansion, any holes in the object expand as well. (The expansion is exaggerated.)

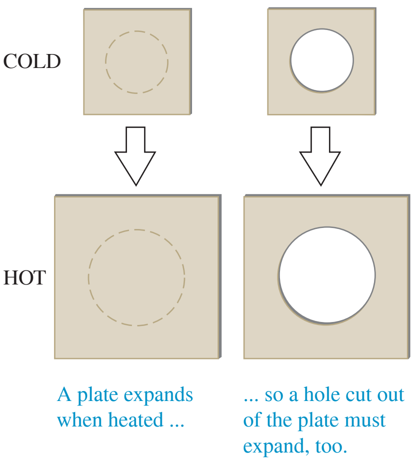
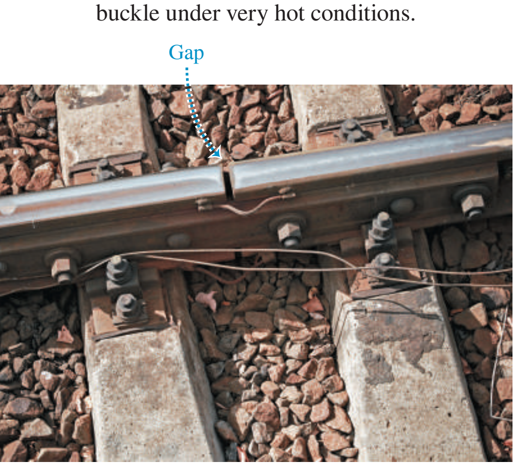


Figure 17.11 This railroad track has a gap between segments to allow for thermal expansion. (The "clickety-clack" sound familiar to railroad passengers comes from the wheels passing over such gaps.) On hot days, the segments expand and fill in the gap. If there were no gaps, the track could buckle under very hot conditions.



CAUTION **Heating an object with a hole** If a solid object has a hole in it, what happens to the size of the hole when the temperature of the object increases? A common misconception is that if the object expands, the hole will shrink because material expands into the hole. But, in fact, if the object expands, the hole will expand too (**Fig. 17.10**); *every* linear dimension of an object changes in the same way when the temperature changes. Think of the atoms in Fig. 17.9a as outlining a cubical hole. When the object expands, the atoms move apart and the hole increases in size. The only situation in which a "hole" will fill in due to thermal expansion is when two separate objects expand and close the gap between them (**Fig. 17.11**).

The direct proportionality in Eq. (17.6) is not exact; it is *approximately* correct only for sufficiently small temperature changes. For a given material, α varies somewhat with the initial temperature T_0 and the size of the temperature interval. We'll ignore this complication here, however. **Table 17.1** lists values of α for several materials. Within the precision of these values we don't need to worry whether T_0 is 0°C or 20°C or some other temperature. Typical values of α are very small; even for a temperature change of 100 C°, the fractional length change $\Delta L/L_0$ is only of the order of $\frac{1}{1000}$ for the metals in the table.

Volume Expansion

Increasing temperature usually causes increases in *volume* for both solids and liquids. Just as with linear expansion, experiments show that if the temperature change ΔT is less than 100 C° or so, the increase in volume ΔV is approximately proportional to both the temperature change ΔT and the initial volume V_0:

Volume thermal expansion:

$$\underbrace{\Delta V}_{\text{Change in volume}} = \underbrace{\beta}_{\text{Coefficient of volume expansion}} \overbrace{V_0}^{\text{Original volume}} \underbrace{\Delta T}_{\text{Temperature change}} \qquad (17.8)$$

The constant β characterizes the volume expansion properties of a particular material; it is called the **coefficient of volume expansion.** The units of β are K^{-1} or $(C°)^{-1}$. As with linear expansion, β varies somewhat with temperature, and Eq. (17.8) is an approximate relationship that is valid only for small temperature changes. For many substances, β decreases at low temperatures. **Table 17.2** lists values of β for several materials near room temperature. Note that the values for liquids are generally much larger than those for solids.

For solid materials we can find a simple relationship between the volume expansion coefficient β and the linear expansion coefficient α. Consider a cube of material with side length L and volume $V = L^3$. At the initial temperature the values are L_0 and V_0. When the temperature increases by dT, the side length increases by dL and the volume increases by an amount dV:

$$dV = \frac{dV}{dL}dL = 3L^2\,dL$$

TABLE 17.1 Coefficients of Linear Expansion

Material	α [K^{-1} or $(C°)^{-1}$]
Aluminum	2.4×10^{-5}
Brass	2.0×10^{-5}
Copper	1.7×10^{-5}
Glass	$0.4–0.9 \times 10^{-5}$
Invar (nickel–iron alloy)	0.09×10^{-5}
Quartz (fused)	0.04×10^{-5}
Steel	1.2×10^{-5}

TABLE 17.2 Coefficients of Volume Expansion

Solids	β [K^{-1} or $(C°)^{-1}$]	Liquids	β [K^{-1} or $(C°)^{-1}$]
Aluminum	7.2×10^{-5}	Ethanol	75×10^{-5}
Brass	6.0×10^{-5}	Carbon disulfide	115×10^{-5}
Copper	5.1×10^{-5}	Glycerin	49×10^{-5}
Glass	$1.2–2.7 \times 10^{-5}$	Mercury	18×10^{-5}
Invar	0.27×10^{-5}		
Quartz (fused)	0.12×10^{-5}		
Steel	3.6×10^{-5}		

Now we replace L and V by the initial values L_0 and V_0. From Eq. (17.6), dL is

$$dL = \alpha L_0\, dT$$

Since $V_0 = L_0^3$, this means that dV can also be expressed as

$$dV = 3L_0^2 \alpha L_0\, dT = 3\alpha V_0\, dT$$

This is consistent with the infinitesimal form of Eq. (17.8), $dV = \beta V_0\, dT$, only if

$$\beta = 3\alpha \qquad (17.9)$$

(Check this relationship for some of the materials listed in Tables 17.1 and 17.2.)

PROBLEM-SOLVING STRATEGY 17.1 Thermal Expansion

IDENTIFY *the relevant concepts:* Decide whether the problem involves changes in length (linear thermal expansion) or in volume (volume thermal expansion).

SET UP *the problem* using the following steps:

1. List the known and unknown quantities and identify the target variables.
2. Choose Eq. (17.6) for linear expansion and Eq. (17.8) for volume expansion.

EXECUTE *the solution* as follows:

1. Solve for the target variables. If you are given an initial temperature T_0 and must find a final temperature T corresponding to a given length or volume change, find ΔT and calculate $T = T_0 + \Delta T$. Remember that the size of a hole in a material varies with temperature just as any other linear dimension, and that the volume of a hole (such as the interior of a container) varies just as that of the corresponding solid shape.
2. Maintain unit consistency. Both L_0 and ΔL (or V_0 and ΔV) must have the same units. If you use a value of α or β in K^{-1} or $(C°)^{-1}$, then ΔT must be in either kelvins or Celsius degrees; from Example 17.1, the two scales are equivalent *for temperature differences.*

EVALUATE *your answer:* Check whether your results make sense.

EXAMPLE 17.2 Length change due to temperature change

WITH VARIATION PROBLEMS

A surveyor uses a steel measuring tape that is exactly 50.000 m long at a temperature of 20°C. The markings on the tape are calibrated for this temperature. (a) What is the length of the tape when the temperature is 35°C? (b) When it is 35°C, the surveyor uses the tape to measure a distance. The value that she reads off the tape is 35.794 m. What is the actual distance?

IDENTIFY and SET UP This problem concerns the linear expansion of a measuring tape. We are given the tape's initial length $L_0 = 50.000$ m at $T_0 = 20°C$. In part (a) we use Eq. (17.6) to find the change ΔL in the tape's length at $T = 35°C$, and use Eq. (17.7) to find L. (Table 17.1 gives the value of α for steel.) Since the tape expands, at 35°C the distance between two successive meter marks is greater than 1 m. Hence the actual distance in part (b) is *larger* than the distance read off the tape by a factor equal to the ratio of the tape's length L at 35°C to its length L_0 at 20°C.

EXECUTE (a) The temperature change is $\Delta T = T - T_0 = 15\ C° = 15$ K; from Eqs. (17.6) and (17.7),

$$\Delta L = \alpha L_0\, \Delta T = (1.2 \times 10^{-5}\ K^{-1})(50\,m)(15\ K)$$

$$= 9.0 \times 10^{-3}\ m = 9.0\ mm$$

$$L = L_0 + \Delta L = 50.000\ m + 0.009\ m = 50.009\ m$$

(b) Our result from part (a) shows that at 35°C, the slightly expanded tape reads a distance of 50.000 m when the true distance is 50.009 m. We can rewrite the algebra of part (a) as $L = L_0(1 + \alpha\,\Delta T)$; at 35°C, *any* true distance will be greater than the reading by the factor $50.009/50.000 = 1 + \alpha\,\Delta T = 1 + 1.8 \times 10^{-4}$. The true distance is therefore

$$(1 + 1.8 \times 10^{-4})(35.794\ m) = 35.800\ m$$

EVALUATE In part (a) we needed only two of the five significant figures of L_0 to compute ΔL to the same number of decimal places as L_0. Our result shows that metals expand very little under moderate temperature changes. However, even the small difference 0.009 m = 9 mm found in part (b) between the scale reading and the true distance can be important in precision work.

KEYCONCEPT A change in temperature causes the *length* of an object to change by an amount that is approximately proportional to the object's initial length and to the temperature change ΔT.

EXAMPLE 17.3 Volume change due to temperature change

WITH VARIATION PROBLEMS

A 200 cm^3 glass flask is filled to the brim with mercury at 20°C. How much mercury overflows when the temperature of the system is raised to 100°C? The coefficient of *linear* expansion of the glass is $0.40 \times 10^{-5}\,K^{-1}$.

IDENTIFY and SET UP This problem involves the volume expansion of the glass and of the mercury. The amount of overflow depends on the *difference* between the volume changes ΔV for these two materials, both given by Eq. (17.8). The mercury will overflow if its coefficient of volume expansion β (see Table 17.2) is greater than that of glass, which we find from Eq. (17.9) using the given value of α.

EXECUTE From Table 17.2, $\beta_{Hg} = 18 \times 10^{-5}\,K^{-1}$. That is indeed greater than $\beta_{glass} = 3\alpha_{glass} = 3(0.40 \times 10^{-5}\,K^{-1}) = 1.2 \times 10^{-5}\,K^{-1}$, from Eq. (17.9). The volume overflow is then

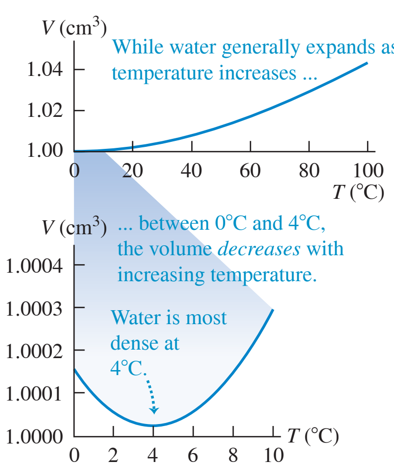
$$\begin{aligned}\Delta V_{Hg} - \Delta V_{glass} &= \beta_{Hg} V_0 \Delta T - \beta_{glass} V_0 \Delta T \\ &= V_0 \Delta T(\beta_{Hg} - \beta_{glass}) \\ &= (200\,cm^3)(80\,C°)(18 \times 10^{-5} - 1.2 \times 10^{-5}) = 2.7\,cm^3\end{aligned}$$

EVALUATE This is basically how a mercury-in-glass thermometer works; the column of mercury inside a sealed tube rises as T increases because mercury expands faster than glass.

As Tables 17.1 and 17.2 show, glass has smaller coefficients of expansion α and β than do most metals. This is why you can use hot water to loosen a metal lid on a glass jar; the metal expands more than the glass does.

KEYCONCEPT A change in temperature causes the *volume* of an object to change by an amount that is approximately proportional to the object's initial volume and to the temperature change ΔT.

Figure 17.12 The volume of 1 gram of water in the temperature range from 0°C to 100°C. By 100°C the volume has increased to 1.043 cm^3. If the coefficient of volume expansion were constant, the curve would be a straight line.



Thermal Expansion of Water

Water, in the temperature range from 0°C to 4°C, *decreases* in volume with increasing temperature. In this range its coefficient of volume expansion is *negative*. Above 4°C, water expands when heated (**Fig. 17.12**). Hence water has its greatest density at 4°C. Water also expands when it freezes, which is why ice humps up in the middle of the compartments in an ice-cube tray. By contrast, most materials contract when they freeze.

This anomalous behavior of water has an important effect on plant and animal life in lakes. A lake cools from the surface down; above 4°C, the cooled water at the surface flows to the bottom because of its greater density. But when the surface temperature drops below 4°C, the water near the surface is less dense than the warmer water below. Hence the downward flow ceases, and the water near the surface remains colder than that at the bottom. As the surface freezes, the ice floats because it is less dense than water. The water at the bottom remains at 4°C until nearly the entire lake is frozen. If water behaved like most substances, contracting continuously on cooling and freezing, lakes would freeze from the bottom up. Circulation due to density differences would continuously carry warmer water to the surface for efficient cooling, and lakes would freeze solid much more easily. This would destroy all plant and animal life that cannot withstand freezing. If water did not have its special properties, the evolution of life would have taken a very different course.

Thermal Stress

If we clamp the ends of a rod rigidly to prevent expansion or contraction and then change the temperature, **thermal stresses** develop. The rod would like to expand or contract, but the clamps won't let it. The resulting stresses may become large enough to strain the rod irreversibly or even break it. (Review the discussion of stress and strain in Section 11.4.)

Engineers must account for thermal stress when designing structures (see Fig. 17.11). Concrete highways and bridge decks usually have gaps between sections, filled with a flexible material or bridged by interlocking teeth (**Fig. 17.13**), to permit expansion and contraction of the concrete. Long steam pipes have expansion joints or U-shaped sections to prevent buckling or stretching with temperature changes. If one end of a steel bridge is rigidly fastened to its abutment, the other end usually rests on rollers.

Figure 17.13 Expansion joints on bridges are needed to accommodate changes in length that result from thermal expansion.

To calculate the thermal stress in a clamped rod, we compute the amount the rod *would* expand (or contract) if not held and then find the stress needed to compress (or stretch) it back to its original length. Suppose that a rod with length L_0 and cross-sectional area A is held at constant length while the temperature is reduced, causing a tensile stress. From Eq. (17.6), the fractional change in length if the rod were free to contract would be

$$\left(\frac{\Delta L}{L_0}\right)_{thermal} = \alpha\,\Delta T \tag{17.10}$$

Since the temperature decreases, both ΔL and ΔT are negative. The tension must increase by an amount F that is just enough to produce an equal and opposite fractional change in length $(\Delta L/L_0)_{\text{tension}}$. From the definition of Young's modulus, Eq. (11.10),

$$Y = \frac{F/A}{\Delta L/L_0} \quad \text{so} \quad \left(\frac{\Delta L}{L_0}\right)_{\text{tension}} = \frac{F}{AY} \tag{17.11}$$

If the length is to be constant, the *total* fractional change in length must be zero. From Eqs. (17.10) and (17.11), this means that

$$\left(\frac{\Delta L}{L_0}\right)_{\text{thermal}} + \left(\frac{\Delta L}{L_0}\right)_{\text{tension}} = \alpha\,\Delta T + \frac{F}{AY} = 0$$

Solve for the tensile stress F/A required to keep the rod's length constant:

Thermal stress: Force needed to keep length of rod constant (F); Cross-sectional area of rod (A); Young's modulus (Y); Coefficient of linear expansion (α); Temperature change (ΔT)

$$\frac{F}{A} = -Y\alpha\,\Delta T \tag{17.12}$$

For a decrease in temperature, ΔT is negative, so F and F/A are positive; this means that a *tensile* force and stress are needed to maintain the length. If ΔT is positive, F and F/A are negative, and the required force and stress are *compressive*.

If there are temperature differences within an object, nonuniform expansion or contraction will result and thermal stresses can be induced. You can break a glass bowl by pouring very hot water into it; the thermal stress between the hot and cold parts of the bowl exceeds the breaking stress of the glass, causing cracks. The same phenomenon makes ice cubes crack when dropped into warm water.

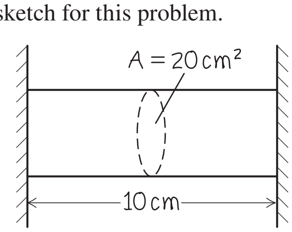
EXAMPLE 17.4 Thermal stress

WITH VARIATION PROBLEMS

An aluminum cylinder 10 cm long, with a cross-sectional area of 20 cm^2, is used as a spacer between two steel walls. At 17.2°C it just slips between the walls. Calculate the stress in the cylinder and the total force it exerts on each wall when it warms to 22.3°C, assuming that the walls are perfectly rigid and a constant distance apart.

IDENTIFY and SET UP See **Fig. 17.14**. The cylinder of given cross-sectional area A exerts force F on each wall; our target variables are the stress F/A and F itself. We use Eq. (17.12) to relate F/A to the temperature change ΔT, and from that calculate F. (The length of the cylinder is irrelevant.) We find Young's modulus Y_{Al} and the coefficient of linear expansion α_{Al} from Tables 11.1 and 17.1, respectively.

Figure **17.14** Our sketch for this problem.



EXECUTE We have $Y_{\text{Al}} = 7.0 \times 10^{10}$ Pa and $\alpha_{\text{Al}} = 2.4 \times 10^{-5}\ \text{K}^{-1}$, and $\Delta T = 22.3°\text{C} - 17.2°\text{C} = 5.1\ \text{C}° = 5.1\ \text{K}$. From Eq. (17.12), the stress is

$$\begin{aligned}\frac{F}{A} &= -Y_{\text{Al}}\alpha_{\text{Al}}\Delta T \\ &= -(7.0 \times 10^{10}\ \text{Pa})(2.4 \times 10^{-5}\ \text{K}^{-1})(5.1\ \text{K}) \\ &= -8.6 \times 10^{6}\ \text{Pa} = -1200\ \text{lb/in.}^2\end{aligned}$$

The total force is the cross-sectional area times the stress:

$$\begin{aligned}F = A\left(\frac{F}{A}\right) &= (20 \times 10^{-4}\ \text{m}^2)(-8.6 \times 10^{6}\ \text{Pa}) \\ &= -1.7 \times 10^{4}\ \text{N} = -1.9\ \text{tons}\end{aligned}$$

EVALUATE The stress on the cylinder and the force it exerts on each wall are immense. Such thermal stresses must be accounted for in engineering.

KEYCONCEPT To keep the length of an object constant when the temperature changes, forces must be applied to both of its ends. The required stress (force per unit area) is proportional to the temperature change.

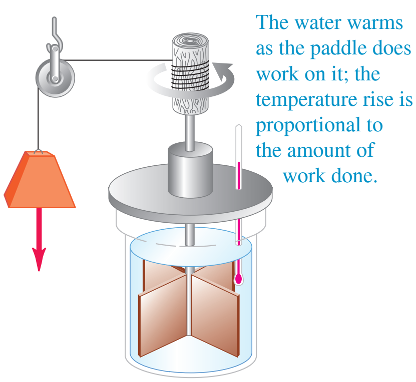
TEST YOUR UNDERSTANDING OF SECTION 17.4 In the bimetallic strip shown in Fig. 17.3a, metal 1 is copper. Which of the following materials could be used for metal 2? (There may be more than one correct answer). (i) Steel; (ii) brass; (iii) aluminum.

ANSWER

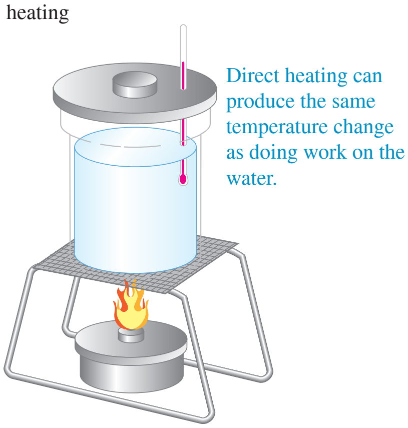
| **(ii) and (iii)** Metal 2 must expand more than metal 1 when heated and so must have a larger coefficient of linear expansion α. From Table 17.1, brass and aluminum have larger values of α than copper, but steel does not.

Figure **17.15** The same temperature change of the same system may be accomplished by (a) doing work on it or (b) adding heat to it.

(a) Raising the temperature of water by doing work on it



(b) Raising the temperature of water by direct heating



17.5 QUANTITY OF HEAT

When you put a cold spoon into a cup of hot coffee, the spoon warms up and the coffee cools down as they approach thermal equilibrium. What causes these temperature changes is a transfer of *energy* from one substance to another. Energy transfer that takes place solely because of a temperature difference is called *heat flow* or *heat transfer,* and energy transferred in this way is called **heat.**

An understanding of the relationship between heat and other forms of energy emerged during the 18th and 19th centuries. Sir James Joule (1818–1889) studied how water can be warmed by vigorous stirring with a paddle wheel (**Fig. 17.15a**). The paddle wheel adds energy to the water by doing *work* on it, and Joule found that *the temperature rise is directly proportional to the amount of work done.* The same temperature change can also be caused by putting the water in contact with some hotter object (Fig. 17.15b); hence this interaction must also involve an energy exchange. We'll explore the relationship between heat and mechanical energy in Chapters 19 and 20.

CAUTION **Temperature vs. heat** It is absolutely essential for you to distinguish between *temperature* and *heat.* Temperature depends on the physical state of a material and is a quantitative description of its hotness or coldness. In physics the term "heat" always refers to energy in transit from one object or system to another because of a temperature difference, never to the amount of energy contained within a particular system. We can change the temperature of an object by adding heat to it or taking heat away, or by adding or subtracting energy in other ways, such as mechanical work (Fig. 17.15a). If we cut an object in half, each half has the same temperature as the whole; but to raise the temperature of each half by a given interval, we add *half* as much heat as for the whole.

We can define a *unit* of quantity of heat based on temperature changes of some specific material. The **calorie** (abbreviated cal) is *the amount of heat required to raise the temperature of 1 gram of water from 14.5°C to 15.5°C.* A food-value calorie is actually a kilocalorie (kcal), equal to 1000 cal. A corresponding unit of heat that uses Fahrenheit degrees and British units is the **British thermal unit,** or Btu. One Btu is the quantity of heat required to raise the temperature of 1 pound (weight) of water 1 F° from 63°F to 64°F.

Because heat is energy in transit, there must be a definite relationship between these units and the familiar mechanical energy units such as the joule (**Fig. 17.16**). Experiments similar in concept to Joule's have shown that

$$\begin{aligned} 1\ \text{cal} &= 4.186\ \text{J} \\ 1\ \text{kcal} &= 1000\ \text{cal} = 4186\ \text{J} \\ 1\ \text{Btu} &= 778\ \text{ft} \cdot \text{lb} = 252\ \text{cal} = 1055\ \text{J} \end{aligned}$$

The calorie is not a fundamental SI unit. The International Committee on Weights and Measures recommends using the joule as the basic unit of energy in all forms, including heat. We'll follow that recommendation in this book.

Figure **17.16** The word "energy" is of Greek origin. This label on a can of Greek coffee shows that 100 milliliters of prepared coffee have an energy content ($\varepsilon\upsilon\acute{\varepsilon}\rho\gamma\varepsilon\iota\alpha$) of 9.6 kilojoules or 2.3 kilocalories.

Specific Heat

We use the symbol Q for quantity of heat. When it is associated with an infinitesimal temperature change dT, we call it dQ. The quantity of heat Q required to increase the temperature of a mass m of a certain material from T_1 to T_2 is found to be approximately proportional to the temperature change $\Delta T = T_2 - T_1$. It is also proportional to the mass m of material. When you're heating water to make tea, you need twice as much heat for two cups as for one if the temperature change is the same. The quantity of heat needed also depends on the nature of the material; raising the temperature of 1 kilogram of water by 1 C° requires 4190 J of heat, but only 910 J is needed to raise the temperature of 1 kilogram of aluminum by 1 C°.

Putting all these relationships together, we have

Heat required to change temperature of a certain mass; Mass of material; Temperature change; Specific heat of material

$$Q = mc\,\Delta T \qquad (17.13)$$

The **specific heat** c has different values for different materials. For an infinitesimal temperature change dT and corresponding quantity of heat dQ,

$$dQ = mc\,dT \tag{17.14}$$

$$c = \frac{1}{m}\frac{dQ}{dT} \qquad \text{(specific heat)} \tag{17.15}$$

In Eqs. (17.13), (17.14), and (17.15), when Q (or dQ) and ΔT (or dT) are positive, heat enters the object and its temperature increases. When they are negative, heat leaves the object and its temperature decreases.

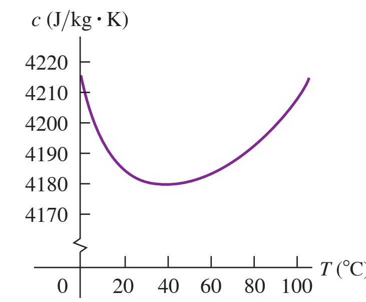
CAUTION **The definition of heat** Remember that dQ does not represent a change in the amount of heat *contained* in an object. Heat is always energy *in transit* as a result of a temperature difference. There is no such thing as "the amount of heat in an object."

The specific heat of water is approximately

$$4190\ \text{J/kg}\cdot\text{K} \qquad 1\ \text{cal/g}\cdot\text{C}^\circ \qquad \text{or} \qquad 1\ \text{Btu/lb}\cdot\text{F}^\circ$$

The specific heat of a material always depends somewhat on the initial temperature and the temperature interval. **Figure 17.17** shows this dependence for water. In this chapter we'll usually ignore this small variation.

Figure 17.17 Specific heat of water as a function of temperature. The value of c varies by less than 1% between 0°C and 100°C.



EXAMPLE 17.5 Feed a cold, starve a fever

During a bout with the flu an 80 kg man ran a fever of 39.0°C (102.2°F) instead of the normal body temperature of 37.0°C (98.6°F). Assuming that the human body is mostly water, how much heat is required to raise his temperature by that amount?

IDENTIFY and SET UP This problem uses the relationship among heat (the target variable), mass, specific heat, and temperature change. We use Eq. (17.13) to determine the required heat Q, with $m = 80$ kg, $c = 4190$ J/kg · K (for water), and $\Delta T = 39.0°\text{C} - 37.0°\text{C} = 2.0\ \text{C}^\circ = 2.0$ K.

EXECUTE From Eq. (17.13),

$$Q = mc\,\Delta T = (80\ \text{kg})(4190\ \text{J/kg}\cdot\text{K})(2.0\ \text{K}) = 6.7 \times 10^5\ \text{J}$$

EVALUATE This corresponds to 160 kcal. In fact, the specific heat of the human body is about 3480 J/kg · K, 83% that of water, because protein, fat, and minerals have lower specific heats. Hence a more accurate answer is $Q = 5.6 \times 10^5$ J = 133 kcal. Either result shows us that were it not for the body's temperature-regulating systems, taking in energy in the form of food would produce measurable changes in body temperature. (The elevated temperature of a person with the flu results from the body's extra activity in response to infection.)

KEYCONCEPT To find the amount of heat required to change the temperature of a mass m of material by an amount ΔT, multiply m by ΔT and by the specific heat c of the material. Heat is positive when it flows into an object and negative when it flows out.

EXAMPLE 17.6 Overheating electronics

You are designing an electronic circuit element made of 23 mg of silicon. The electric current through it adds energy at the rate of 7.4 mW = 7.4×10^{-3} J/s. If your design doesn't allow any heat transfer out of the element, at what rate does its temperature increase? The specific heat of silicon is 705 J/kg · K.

IDENTIFY and SET UP The energy added to the circuit element gives rise to a temperature increase, just as if heat were flowing into the element at the rate $dQ/dt = 7.4 \times 10^{-3}$ J/s. Our target variable is the rate of temperature change dT/dt. We can use Eq. (17.14),which relates infinitesimal temperature changes dT to the corresponding heat dQ, to obtain an expression for dQ/dt in terms of dT/dt.

EXECUTE We divide both sides of Eq. (17.14) by dt and rearrange:

$$\frac{dT}{dt} = \frac{dQ/dt}{mc} = \frac{7.4 \times 10^{-3}\ \text{J/s}}{(23 \times 10^{-6}\ \text{kg})(705\ \text{J/kg}\cdot\text{K})} = 0.46\ \text{K/s}$$

EVALUATE At this rate of temperature rise (27 K/min), the circuit element would soon self-destruct. Heat transfer is an important design consideration in electronic circuit elements.

KEYCONCEPT Any energy flow (not just heat) into or out of a quantity of material can cause the temperature of the material to change. The rate of energy flow is equal to the mass times the specific heat of the material times the rate of temperature change.

Molar Heat Capacity

Sometimes it's more convenient to describe a quantity of substance in terms of the number of *moles* n rather than the *mass* m of material. Recall from your study of chemistry that a mole of any pure substance always contains the same number of molecules. (We'll discuss this point in more detail in Chapter 18.) The *molar mass* of any substance, denoted by M, is the mass per mole. (The quantity M is sometimes called *molecular weight,* but *molar mass* is preferable; the quantity depends on the mass of a molecule, not its weight.) For example, the molar mass of water is $18.0\ \text{g/mol} = 18.0 \times 10^{-3}\ \text{kg/mol}$; 1 mole of water has a mass of $18.0\ \text{g} = 0.0180\ \text{kg}$. The total mass m of material is equal to the mass per mole M times the number of moles n:

$$m = nM \tag{17.16}$$

Replacing the mass m in Eq. (17.13) by the product nM, we find

$$Q = nMc\,\Delta T \tag{17.17}$$

The product Mc is called the **molar heat capacity** (or *molar specific heat*) and is denoted by C (capitalized). With this notation we rewrite Eq. (17.17) as

$$\underbrace{Q}_{\substack{\text{Heat required to}\\ \text{change temperature}\\ \text{of a certain}\\ \text{number of moles}}} = \overbrace{n}^{\text{Number of moles of material}}\underbrace{C}_{\text{Molar heat capacity of material}}\overbrace{\Delta T}^{\text{Temperature change}} \tag{17.18}$$

Figure **17.18** Water has a much higher specific heat than the glass or metals used to make cookware. This helps explain why it takes several minutes to boil water on a stove, even though the pot or kettle reaches a high temperature very quickly.

Comparing to Eq. (17.15), we can express the molar heat capacity C (heat per mole per temperature change) in terms of the specific heat c (heat per mass per temperature change) and the molar mass M (mass per mole):

$$C = \frac{1}{n}\frac{dQ}{dT} = Mc \qquad \text{(molar heat capacity)} \tag{17.19}$$

For example, the molar heat capacity of water is

$$\begin{aligned} C = Mc &= (0.0180\ \text{kg/mol})(4190\ \text{J/kg}\cdot\text{K}) \\ &= 75.4\ \text{J/mol}\cdot\text{K} \end{aligned}$$

Table 17.3 gives values of specific heat and molar heat capacity for several substances. Note the remarkably large specific heat for water (**Fig. 17.18**).

CAUTION **The meaning of "heat capacity"** The term "heat capacity" is unfortunate because it gives the erroneous impression that an object *contains* a certain amount of heat. Remember, heat is energy in transit to or from an object, not the energy residing in the object.

Measurements of specific heats and molar heat capacities for solid materials are usually made at constant atmospheric pressure; the corresponding values are called the *specific heat* and *molar heat capacity at constant pressure,* denoted by c_p and C_p. For a gas it is usually easier to keep the substance in a container with constant *volume;* the corresponding values are called the *specific heat* and *molar heat capacity at constant volume,* denoted by c_V and C_V. For a given substance, C_V and C_p are different. If the system can expand while heat is added, there is additional energy exchange through the performance of *work* by the system on its surroundings. If the volume is constant, the system does no work. For gases the difference between C_p and C_V is substantial. We'll study heat capacities of gases in detail in Section 19.7.

TABLE 17.3 Approximate Specific Heats and Molar Heat Capacities (Constant Pressure)

Substance	Specific Heat, c (J/kg · K)	Molar Mass, M (kg/mol)	Molar Heat Capacity, C (J/mol · K)
Aluminum	910	0.0270	24.6
Beryllium	1970	0.00901	17.7
Copper	390	0.0635	24.8
Ethanol	2428	0.0461	111.9
Ethylene glycol	2386	0.0620	148.0
Ice (near 0°C)	2100	0.0180	37.8
Iron	470	0.0559	26.3
Lead	130	0.207	26.9
Marble ($CaCO_3$)	879	0.100	87.9
Mercury	138	0.201	27.7
Salt (NaCl)	879	0.0585	51.4
Silver	234	0.108	25.3
Water (liquid)	4190	0.0180	75.4

The last column of Table 17.3 shows something interesting. The molar heat capacities for most elemental solids are about the same: about 25 J/mol · K. This correlation, named the *rule of Dulong and Petit* (for its discoverers), forms the basis for a very important idea. The number of atoms in 1 mole is the same for all elemental substances. This means that on a *per atom* basis, about the same amount of heat is required to raise the temperature of each of these elements by a given amount, even though the *masses* of the atoms are very different. The heat required for a given temperature increase depends only on *how many* atoms the sample contains, not on the mass of an individual atom. We'll see the reason the rule of Dulong and Petit works so well when we study the molecular basis of heat capacities in greater detail in Chapter 18.

TEST YOUR UNDERSTANDING OF SECTION 17.5 You wish to raise the temperature of each of the following samples from 20°C to 21°C. Rank these in order of the amount of heat needed to do this, from highest to lowest. (i) 1 kilogram of mercury; (ii) 1 kilogram of ethanol; (iii) 1 mole of mercury; (iv) 1 mole of ethanol.

ANSWER

(ii), (i), (iv), (iii) For (i) and (ii), the relevant quantity is the specific heat c of the substance, which is the amount of heat required to raise the temperature of 1 *kilogram* of that substance by 1 K (1 C°). From Table 17.3, these values are (i) 138 J for mercury and (ii) 2428 J for ethanol. For (iii) and (iv) we need the molar heat capacity C, which is the amount of heat required to raise the temperature of 1 *mole* of that substance by 1 C°. Again from Table 17.3, these values are (iii) 27.7 J for mercury and (iv) 111.9 J for ethanol. (The ratio of molar heat capacities is different from the ratio of the specific heats because a mole of mercury and a mole of ethanol have different masses.)

17.6 CALORIMETRY AND PHASE CHANGES

Calorimetry means "measuring heat." We have discussed the energy transfer (heat) involved in temperature changes. Heat is also involved in *phase changes,* such as the melting of ice or boiling of water. Once we understand these additional heat relationships, we can analyze a variety of problems involving quantity of heat.

Phase Changes

We use the term **phase** to describe a specific state of matter, such as a solid, liquid, or gas. The compound H_2O exists in the *solid phase* as ice, in the *liquid phase* as water, and in the *gaseous phase* as steam. (These are also referred to as **states of matter:** the solid state, the liquid state, and the gaseous state.) A transition from one phase to another is called a **phase change** or *phase transition*. For any given pressure a phase change takes place at a definite temperature, usually accompanied by heat flowing in or out and a change of volume and density.

Figure 17.19 The surrounding air is at room temperature, but this ice–water mixture remains at 0°C until all of the ice has melted and the phase change is complete.

A familiar phase change is the melting of ice. When we add heat to ice at 0°C and normal atmospheric pressure, the temperature of the ice *does not* increase. Instead, some of it melts to form liquid water. If we add the heat slowly, to maintain the system very close to thermal equilibrium, the temperature remains at 0°C until all the ice is melted (**Fig. 17.19**). The effect of adding heat to this system is not to raise its temperature but to change its *phase* from solid to liquid.

To change 1 kg of ice at 0°C to 1 kg of liquid water at 0°C and normal atmospheric pressure requires 3.34×10^5 J of heat. The heat required per unit mass is called the **heat of fusion** (or sometimes *latent heat of fusion*), denoted by L_f. For water at normal atmospheric pressure the heat of fusion is

$$L_f = 3.34 \times 10^5 \text{ J/kg} = 79.6 \text{ cal/g} = 143 \text{ Btu/lb}$$

More generally, to melt a mass m of material that has a heat of fusion L_f requires a quantity of heat Q given by

$$Q = mL_f$$

This process is *reversible*. To freeze liquid water to ice at 0°C, we have to *remove* heat; the magnitude is the same, but in this case, Q is negative because heat is removed rather than added. To cover both possibilities and to include other kinds of phase changes, we write

Heat transfer in a phase change

Mass of material that changes phase

$$Q = \pm mL \qquad (17.20)$$

Latent heat for this phase change

+ if heat enters material, − if heat leaves

The plus sign (heat entering) is used when the material melts; the minus sign (heat leaving) is used when it freezes. The heat of fusion is different for different materials, and it also varies somewhat with pressure.

For any given material at any given pressure, the freezing temperature is the same as the melting temperature. At this unique temperature the liquid and solid phases can coexist in a condition called **phase equilibrium.**

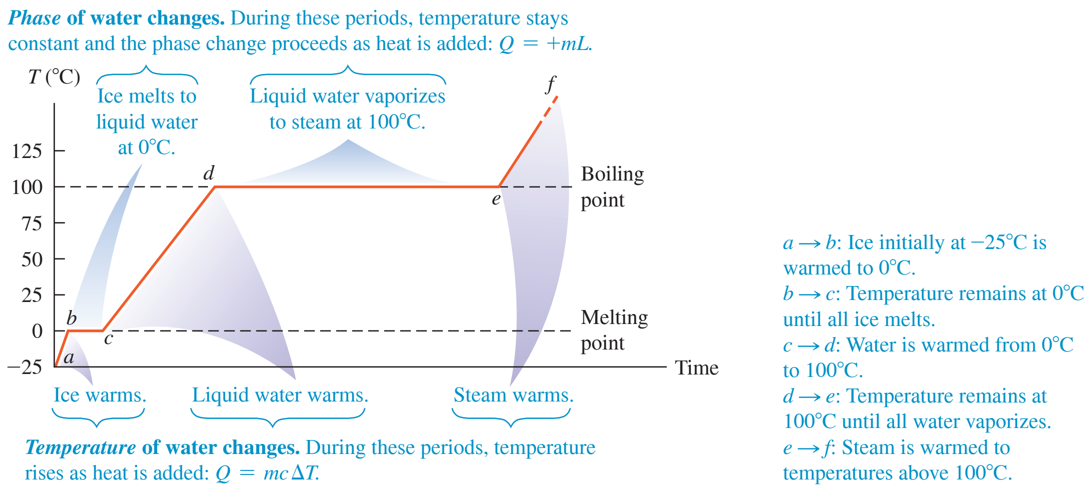
We can go through this whole story again for *boiling* or *evaporation,* a phase transition between liquid and gaseous phases. The corresponding heat (per unit mass) is called the **heat of vaporization** L_v. At normal atmospheric pressure the heat of vaporization L_v for water is

$$L_v = 2.256 \times 10^6 \text{ J/kg} = 539 \text{ cal/g} = 970 \text{ Btu/lb}$$

That is, it takes 2.256×10^6 J to change 1 kg of liquid water at 100°C to 1 kg of water vapor at 100°C. By comparison, to raise the temperature of 1 kg of water from 0°C to 100°C requires $Q = mc\,\Delta T = (1.00\text{ kg})(4190\text{ J/kg}\cdot\text{C}°)(100\text{ C}°) = 4.19 \times 10^5$ J, less than one-fifth as much heat as is required for vaporization at 100°C. This agrees with everyday kitchen experience; a pot of water may reach boiling temperature in a few minutes, but it takes a much longer time to completely evaporate all the water away.

Like melting, boiling is a reversible transition. When heat is removed from a gas at the boiling temperature, the gas returns to the liquid phase, or *condenses,* giving up to its surroundings the same quantity of heat (heat of vaporization) that was needed to vaporize it. At a given pressure the boiling and condensation temperatures are always the same; at this temperature the liquid and gaseous phases can coexist in phase equilibrium.

Figure **17.20** Graph of temperature versus time for a specimen of water initially in the solid phase (ice). Heat is added to the specimen at a constant rate. The temperature remains constant during each change of phase, provided that the pressure remains constant.



Both L_v and the boiling temperature of a material depend on pressure. Water boils at a lower temperature (about 95°C) in Denver than in Pittsburgh because Denver is at higher elevation and the average atmospheric pressure is lower. The heat of vaporization is somewhat greater at this lower pressure, about 2.27×10^6 J/kg.

Figure 17.20 summarizes these ideas about phase changes. **Table 17.4** lists heats of fusion and vaporization for some materials and their melting and boiling temperatures at normal atmospheric pressure. Very few *elements* have melting temperatures in the vicinity of ordinary room temperatures; one of the few is the metal gallium, shown in **Fig. 17.21**.

Figure **17.21** The metal gallium, shown here melting in a person's hand, is one of the few elements that melt in the vicinity of room temperature. Its melting temperature is 29.8°C, and its heat of fusion is 8.04×10^4 J/kg.

TABLE 17.4 Heats of Fusion and Vaporization

	Normal Melting Point		Heat of Fusion, L_f (J/kg)	Normal Boiling Point		Heat of Vaporization, L_v (J/kg)
Substance	**K**	**°C**		**K**	**°C**	
Helium	*	*	*	4.216	−268.93	20.9×10^3
Hydrogen	13.84	−259.31	58.6×10^3	20.26	−252.89	452×10^3
Nitrogen	63.18	−209.97	25.5×10^3	77.34	−195.8	201×10^3
Oxygen	54.36	−218.79	13.8×10^3	90.18	−183.0	213×10^3
Ethanol	159	−114	104.2×10^3	351	78	854×10^3
Mercury	234	−39	11.8×10^3	630	357	272×10^3
Water	273.15	0.00	334×10^3	373.15	100.00	2256×10^3
Sulfur	392	119	38.1×10^3	717.75	444.60	326×10^3
Lead	600.5	327.3	24.5×10^3	2023	1750	871×10^3
Antimony	903.65	630.50	165×10^3	1713	1440	561×10^3
Silver	1233.95	960.80	88.3×10^3	2466	2193	2336×10^3
Gold	1336.15	1063.00	64.5×10^3	2933	2660	1578×10^3
Copper	1356	1083	134×10^3	1460	1187	5069×10^3

*A pressure in excess of 25 atmospheres is required to make helium solidify. At 1 atmosphere pressure, helium remains a liquid down to absolute zero.

Figure 17.22 When this airplane flew into a cloud at a temperature just below freezing, the plane struck supercooled water droplets in the cloud that rapidly crystallized and formed ice on the plane's nose (shown here) and wings. Such in-flight icing can be extremely hazardous, which is why commercial airliners are equipped with devices to remove ice.

A substance can sometimes change directly from the solid to the gaseous phase. This process is called *sublimation,* and the solid is said to *sublime.* The corresponding heat is called the *heat of sublimation,* L_s. Liquid carbon dioxide cannot exist at a pressure lower than about 5×10^5 Pa (about 5 atm), and "dry ice" (solid carbon dioxide) sublimes at atmospheric pressure. Sublimation of water from frozen food causes freezer burn. The reverse process, a phase change from gas to solid, occurs when frost forms on cold objects such as refrigerator cooling coils.

Very pure water can be cooled several degrees below the freezing temperature without freezing; the resulting unstable state is described as *supercooled.* When a small ice crystal is dropped in or the water is agitated, it crystallizes within a second or less (**Fig. 17.22**). Supercooled water *vapor* condenses quickly into fog droplets when a disturbance, such as dust particles or ionizing radiation, is introduced. This principle is used in "seeding" clouds, which often contain supercooled water vapor, to cause condensation and rain.

A liquid can sometimes be *superheated* above its normal boiling temperature. Any small disturbance such as agitation causes local boiling with bubble formation.

Steam heating systems for buildings use a boiling–condensing process to transfer heat from the furnace to the radiators. Each kilogram of water that is turned to steam in the boiler absorbs over 2×10^6 J (the heat of vaporization L_v of water) from the boiler and gives it up when it condenses in the radiators. Boiling–condensing processes are also used in refrigerators, air conditioners, and heat pumps. We'll discuss these systems in Chapter 20.

The temperature-control mechanisms of many warm-blooded animals make use of heat of vaporization, removing heat from the body by using it to evaporate water from the tongue (panting) or from the skin (sweating). Such *evaporative cooling* enables humans to maintain normal body temperature in hot, dry desert climates where the air temperature may reach 55°C (about 130°F). The skin temperature may be as much as 30°C cooler than the surrounding air. Under these conditions a normal person may perspire several liters per day, and this lost water must be replaced. Evaporative cooling also explains why you feel cold when you first step out of a swimming pool (**Fig. 17.23**).

Figure 17.23 The water may be warm and it may be a hot day, but these children will feel cold when they first step out of the swimming pool. That's because as water evaporates from their skin, it removes the heat of vaporization from their bodies. To stay warm, they will need to dry off immediately.

Evaporative cooling is also used to condense and recirculate "used" steam in coal-fired or nuclear-powered electric-generating plants. That's what goes on in the large, tapered concrete towers that you see at such plants.

Chemical reactions such as combustion are analogous to phase changes in that they involve definite quantities of heat. Complete combustion of 1 gram of gasoline produces about 46,000 J or about 11,000 cal, so the **heat of combustion** L_c of gasoline is

$$L_c = 46{,}000 \text{ J/g} = 4.6 \times 10^7 \text{ J/kg}$$

Energy values of foods are defined similarly. When we say that a gram of peanut butter "contains 6 calories," we mean that 6 kcal of heat (6000 cal or 25,000 J) is released when the carbon and hydrogen atoms in the peanut butter react with oxygen (with the help of enzymes) and are completely converted to CO_2 and H_2O. Not all of this energy is directly useful for mechanical work. We'll study the *efficiency* of energy utilization in Chapter 20.

Heat Calculations

Let's look at some examples of calorimetry calculations (calculations with heat). The basic principle is very simple: When heat flow occurs between two objects that are isolated from their surroundings, the amount of heat lost by one object must equal the amount gained by the other. Heat is energy in transit, so this principle is really just conservation of energy. Calorimetry, dealing entirely with one conserved quantity, is in many ways the simplest of all physical theories!

PROBLEM-SOLVING STRATEGY 17.2 Calorimetry Problems

IDENTIFY *the relevant concepts:* When heat flow occurs between two or more objects that are isolated from their surroundings, the *algebraic sum* of the quantities of heat transferred to all the objects is zero. We take a quantity of heat *added* to an object as *positive* and a quantity *leaving* an object as *negative.*

SET UP *the problem* using the following steps:

1. Identify the objects that exchange heat.
2. Each object may undergo a temperature change only, a phase change at constant temperature, or both. Use Eq. (17.13) for the heat transferred in a temperature change and Eq. (17.20) for the heat transferred in a phase change.
3. Consult Table 17.3 for values of specific heat or molar heat capacity and Table 17.4 for heats of fusion or vaporization.
4. List the known and unknown quantities and identify the target variables.

EXECUTE *the solution* as follows:

1. Use Eq. (17.13) and/or Eq. (17.20) and the energy-conservation relationship $\sum Q = 0$ to solve for the target variables. Ensure that you use the correct algebraic signs for Q and ΔT terms, and that you correctly write $\Delta T = T_{\text{final}} - T_{\text{initial}}$ and not the reverse.
2. If a phase change occurs, you may not know in advance whether all, or only part, of the material undergoes a phase change. Make a reasonable guess; if that leads to an unreasonable result (such as a final temperature higher or lower than any initial temperature), the guess was wrong. Try again!

EVALUATE *your answer:* Double-check your calculations, and ensure that the results are physically sensible.

EXAMPLE 17.7 A temperature change with no phase change

WITH VARIATION PROBLEMS

A camper pours 0.300 kg of coffee, initially in a pot at 70.0°C, into a 0.120 kg aluminum cup initially at 20.0°C. What is the equilibrium temperature? Assume that coffee has the same specific heat as water and that no heat is exchanged with the surroundings.

IDENTIFY and SET UP The target variable is the common final temperature T of the cup and coffee. No phase changes occur, so we need only Eq. (17.13). With subscripts C for coffee, W for water, and Al for aluminum, we have $T_{0\text{C}} = 70.0°\text{C}$ and $T_{0\text{Al}} = 20.0°\text{C}$; Table 17.3 gives $c_\text{W} = 4190\ \text{J/kg}\cdot\text{K}$ and $c_\text{Al} = 910\ \text{J/kg}\cdot\text{K}$.

EXECUTE The (negative) heat gained by the coffee is $Q_\text{C} = m_\text{C}c_\text{W}\,\Delta T_\text{C}$. The (positive) heat gained by the cup is $Q_\text{Al} = m_\text{Al}c_\text{Al}\,\Delta T_\text{Al}$. We set $Q_\text{C} + Q_\text{Al} = 0$ (see Problem-Solving Strategy 17.2) and substitute $\Delta T_\text{C} = T - T_{0\text{C}}$ and $\Delta T_\text{Al} = T - T_{0\text{Al}}$:

$$Q_\text{C} + Q_\text{Al} = m_\text{C}c_\text{W}\,\Delta T_\text{C} + m_\text{Al}c_\text{Al}\,\Delta T_\text{Al} = 0$$
$$m_\text{C}c_\text{W}(T - T_{0\text{C}}) + m_\text{Al}c_\text{Al}(T - T_{0\text{Al}}) = 0$$

Then we solve this expression for the final temperature T. A little algebra gives

$$T = \frac{m_\text{C}c_\text{W}T_{0\text{C}} + m_\text{Al}c_\text{Al}T_{0\text{Al}}}{m_\text{C}c_\text{W} + m_\text{Al}c_\text{Al}} = 66.0°\text{C}$$

EVALUATE The final temperature is much closer to the initial temperature of the coffee than to that of the cup; water has a much higher specific heat than aluminum, and we have more than twice as much mass of water. We can also find the quantities of heat by substituting the value $T = 66.0°\text{C}$ back into the original equations. We find $Q_\text{C} = -5.0 \times 10^3\ \text{J}$ and $Q_\text{Al} = +5.0 \times 10^3\ \text{J}$. As expected, Q_C is negative: The coffee loses heat to the cup.

KEYCONCEPT In a calorimetry problem in which two objects at different temperatures interact by exchanging heat, energy is conserved: The sum of the heat flows (one positive, one negative) into the two objects is zero. The heat flow stops when the two objects reach the same temperature.

EXAMPLE 17.8 Changes in both temperature and phase

WITH VARIATION PROBLEMS

A glass contains 0.25 kg of Omni-Cola (mostly water) initially at 25°C. How much ice, initially at −20°C, must you add to obtain a final temperature of 0°C with all the ice melted? Ignore the heat capacity of the glass.

IDENTIFY and SET UP The Omni-Cola and ice exchange heat. The cola undergoes a temperature change; the ice undergoes both a temperature change and a phase change from solid to liquid. We use subscripts C for cola, I for ice, and W for water. The target variable is the mass of ice, m_I. We use Eq. (17.13) to obtain an expression for the amount of heat involved in cooling the drink to $T = 0°\text{C}$ and warming the ice to $T = 0°\text{C}$, and Eq. (17.20) to obtain an expression for the heat required to melt the ice at 0°C. We have $T_{0\text{C}} = 25°\text{C}$ and $T_{0\text{I}} = -20°\text{C}$, Table 17.3 gives $c_\text{W} = 4190\ \text{J/kg}\cdot\text{K}$ and $c_\text{I} = 2100\ \text{J/kg}\cdot\text{K}$, and Table 17.4 gives $L_\text{f} = 3.34 \times 10^5\ \text{J/kg}$.

EXECUTE From Eq. (17.13), the (negative) heat gained by the Omni-Cola is $Q_\text{C} = m_\text{C}c_\text{W}\,\Delta T_\text{C}$. The (positive) heat gained by the ice in warming is $Q_\text{I} = m_\text{I}c_\text{I}\,\Delta T_\text{I}$. The (positive) heat required to melt the ice is $Q_2 = m_\text{I}L_\text{f}$. We set $Q_\text{C} + Q_\text{I} + Q_2 = 0$, insert $\Delta T_\text{C} = T - T_{0\text{C}}$ and $\Delta T_\text{I} = T - T_{0\text{I}}$, and solve for m_I:

$$m_\text{C}c_\text{W}\,\Delta T_\text{C} + m_\text{I}c_\text{I}\,\Delta T_\text{I} + m_\text{I}L_\text{f} = 0$$
$$m_\text{C}c_\text{W}(T - T_{0\text{C}}) + m_\text{I}c_\text{I}(T - T_{0\text{I}}) + m_\text{I}L_\text{f} = 0$$
$$m_\text{I}[c_\text{I}(T - T_{0\text{I}}) + L_\text{f}] = -m_\text{C}c_\text{W}(T - T_{0\text{C}})$$
$$m_\text{I} = m_\text{C}\frac{c_\text{W}(T_{0\text{C}} - T)}{c_\text{I}(T - T_{0\text{I}}) + L_\text{f}}$$

Substituting numerical values, we find that $m_\text{I} = 0.070\ \text{kg} = 70\ \text{g}$.

EVALUATE Three or four medium-size ice cubes would make about 70 g, which seems reasonable given the 250 g of Omni-Cola to be cooled.

KEYCONCEPT When heat flows between two objects and one or both of them change phase, your calculations must include the heat required to cause the phase change. This depends on the object's mass and material and on which phase change occurs.

EXAMPLE 17.9 What's cooking?

WITH VARIATION PROBLEMS

A hot copper pot of mass 2.0 kg (including its copper lid) is at a temperature of 150°C. You pour 0.10 kg of cool water at 25°C into the pot, then quickly replace the lid so no steam can escape. Find the final temperature of the pot and its contents, and determine the phase of the water (liquid, gas, or a mixture). Assume that no heat is lost to the surroundings.

IDENTIFY and SET UP The water and the pot exchange heat. Three outcomes are possible: (1) No water boils, and the final temperature T is less than 100°C; (2) some water boils, giving a mixture of water and steam at 100°C; or (3) all the water boils, giving 0.10 kg of steam at 100°C or greater. We use Eq. (17.13) for the heat transferred in a temperature change and Eq. (17.20) for the heat transferred in a phase change.

EXECUTE First consider case (1), which parallels Example 17.8 exactly. The equation that states that the heat flow into the water equals the heat flow out of the pot is

$$Q_W + Q_{Cu} = m_W c_W (T - T_{0W}) + m_{Cu} c_{Cu} (T - T_{0Cu}) = 0$$

Here we use subscripts W for water and Cu for copper, with $m_W = 0.10$ kg, $m_{Cu} = 2.0$ kg, $T_{0W} = 25°C$, and $T_{0Cu} = 150°C$. From Table 17.3, $c_W = 4190$ J/kg·K and $c_{Cu} = 390$ J/kg·K. Solving for the final temperature T and substituting these values, we get

$$T = \frac{m_W c_W T_{0W} + m_{Cu} c_{Cu} T_{0Cu}}{m_W c_W + m_{Cu} c_{Cu}} = 106°C$$

But this is above the boiling point of water, which contradicts our assumption that no water boils! So at least some of the water boils.

So consider case (2), in which the final temperature is $T = 100°C$ and some unknown fraction x of the water boils, where (if this case is correct) x is greater than zero and less than or equal to 1. The (positive) amount of heat needed to vaporize this water is $x m_W L_v$. The energy-conservation condition $Q_W + Q_{Cu} = 0$ is then

$$m_W c_W (100°C - T_{0W}) + x m_W L_v + m_{Cu} c_{Cu} (100°C - T_{0Cu}) = 0$$

We solve for the target variable x:

$$x = \frac{-m_{Cu} c_{Cu} (100°C - T_{0Cu}) - m_W c_W (100°C - T_{0W})}{m_W L_v}$$

With $L_v = 2.256 \times 10^6$ J from Table 17.4, this yields $x = 0.034$. We conclude that the final temperature of the water and copper is 100°C and that $0.034(0.10 \text{ kg}) = 0.0034 \text{ kg} = 3.4$ g of the water is converted to steam at 100°C.

EVALUATE Had x turned out to be greater than 1, case (3) would have held; all the water would have vaporized, and the final temperature would have been greater than 100°C. Can you show that this would have been the case if we had originally poured less than 15 g of 25°C water into the pot?

KEYCONCEPT In many calorimetry problems you won't know whether or not a phase change occurs. To find out, try working the problem three ways (assuming no phase change, assuming part of the object changes phase, and assuming all of it changes phase). The way that leads to a sensible result is the correct one.

EXAMPLE 17.10 Combustion, temperature change, and phase change

In a particular camp stove, only 30% of the energy released in burning gasoline goes to heating the water in a pot on the stove. How much gasoline must we burn to heat 1.00 L (1.00 kg) of water from 20°C to 100°C and boil away 0.25 kg of it?

IDENTIFY and SET UP All of the water undergoes a temperature change and part of it undergoes a phase change, from liquid to gas. We determine the heat required to cause both of these changes, and then use the 30% combustion efficiency to determine the amount of gasoline that must be burned (the target variable). We use Eqs. (17.13) and (17.20) and the idea of heat of combustion.

EXECUTE To raise the temperature of the water from 20°C to 100°C requires

$$Q_1 = mc\,\Delta T = (1.00 \text{ kg})(4190 \text{ J/kg}\cdot\text{K})(80 \text{ K}) = 3.35 \times 10^5 \text{ J}$$

To boil 0.25 kg of water at 100°C requires

$$Q_2 = mL_v = (0.25 \text{ kg})(2.256 \times 10^6 \text{ J/kg}) = 5.64 \times 10^5 \text{ J}$$

The total energy needed is $Q_1 + Q_2 = 8.99 \times 10^5$ J. This is $30\% = 0.30$ of the total heat of combustion, which is therefore $(8.99 \times 10^5 \text{ J})/0.30 = 3.00 \times 10^6$ J. As we mentioned earlier, the combustion of 1 g of gasoline releases 46,000 J, so the mass of gasoline required is $(3.00 \times 10^6 \text{ J})/(46{,}000 \text{ J/g}) = 65$ g, or a volume of about 0.09 L of gasoline.

EVALUATE This result suggests the tremendous amount of energy released in burning even a small quantity of gasoline. Another 123 g of gasoline would be required to boil away the remaining water; can you prove this?

KEYCONCEPT To find the amount of heat released when a quantity of substance undergoes combustion, multiply the mass that combusts by the heat of combustion of the substance.

TEST YOUR UNDERSTANDING OF SECTION 17.6 You take a block of ice at 0°C and add heat to it at a steady rate. It takes a time t to completely convert the block of ice to steam at 100°C. What do you have at time $t/2$? (i) All ice at 0°C; (ii) a mixture of ice and water at 0°C; (iii) water at a temperature between 0°C and 100°C; (iv) a mixture of water and steam at 100°C.

ANSWER

(iv) In time t the system goes from point b to point e in Fig. 17.20. According to this figure, at time $t/2$ (halfway along the horizontal axis from b to e), the system is at 100°C and is still boiling; that is, it is a mixture of liquid and gas. This says that most of the heat added goes into boiling the water.

17.7 MECHANISMS OF HEAT TRANSFER

We have talked about *conductors* and *insulators,* materials that permit or prevent heat transfer between objects. Now let's look in more detail at *rates* of energy transfer. In the kitchen you use a metal or glass pot for good heat transfer from the stove to whatever you're cooking, but your refrigerator is insulated with a material that *prevents* heat from flowing into the food inside the refrigerator. How do we describe the difference between these two materials?

The three mechanisms of heat transfer are conduction, convection, and radiation. *Conduction* occurs within an object or between two objects in contact. *Convection* depends on motion of mass from one region of space to another. *Radiation* is heat transfer by electromagnetic radiation, such as sunshine, with no need for matter to be present in the space between objects.

Conduction

If you hold one end of a copper rod and place the other end in a flame, the end you are holding gets hotter and hotter, even though it is not in direct contact with the flame. Heat reaches the cooler end by **conduction** through the material. The atoms in the hotter regions have more kinetic energy, on the average, than their cooler neighbors. They jostle their neighbors, giving them some of their energy. The neighbors jostle *their* neighbors, and so on through the material. The atoms don't move from one region of material to another, but their energy does.

Most metals also conduct heat by another, more effective mechanism. Within the metal, some electrons can leave their parent atoms and wander through the metal. These "free" electrons can rapidly carry energy from hotter to cooler regions of the metal, so metals are generally good conductors of heat. A metal rod at 20°C feels colder than a piece of wood at 20°C because heat can flow more easily from your hand into the metal. The presence of "free" electrons also causes most metals to be good electrical conductors.

In conduction, the direction of heat flow is always from higher to lower temperature. **Figure 17.24a** shows a rod of conducting material with cross-sectional area A and length L. The left end of the rod is kept at a temperature T_H and the right end at a lower temperature T_C, so heat flows from left to right. The sides of the rod are covered by an ideal insulator, so no heat transfer occurs at the sides.

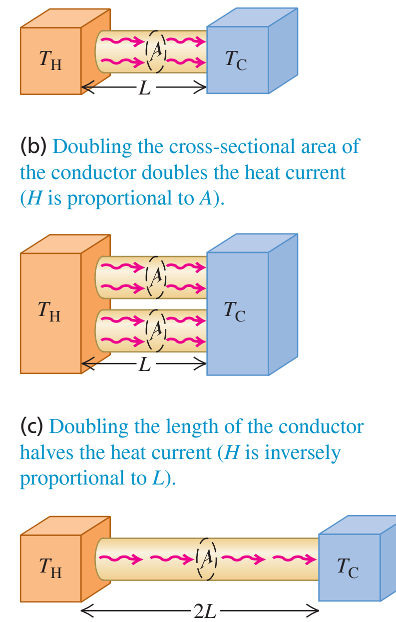
When a quantity of heat dQ is transferred through the rod in a time dt, the rate of heat flow is dQ/dt. We call this rate the **heat current,** denoted by H. That is, $H = dQ/dt$. Experiments show that the heat current is proportional to the cross-sectional area A of the rod (Fig. 17.24b) and to the temperature difference $(T_H - T_C)$ and is inversely proportional to the rod length L (Fig. 17.24c):

$$\underset{\text{Heat current in conduction}}{H} = \overset{\text{Rate of heat flow}}{\frac{dQ}{dt}} = \underset{\text{Thermal conductivity of rod material}}{k}\,\underset{\text{Cross-sectional area of rod}}{A}\,\frac{\overset{\text{Temperatures of hot and cold ends of rod}}{T_H - T_C}}{\underset{\text{Length of rod}}{L}} \qquad (17.21)$$

The quantity $(T_H - T_C)/L$ is the temperature difference *per unit length;* it is called the magnitude of the **temperature gradient.** The numerical value of the **thermal conductivity** k depends on the material of the rod. Materials with large k are good conductors of heat; materials with small k are poor conductors, or insulators. Equation (17.21) also gives the heat current through a slab or through *any* homogeneous object with uniform cross section A perpendicular to the direction of flow; L is the length of the heat-flow path.

The units of heat current H are units of energy per time, or power; the SI unit of heat current is the watt $(1\text{ W} = 1\text{ J/s})$. We can find the units of k by solving Eq. (17.21) for k; you can show that the SI units are $\text{W/m}\cdot\text{K}$. **Table 17.5** gives some numerical values of k.

Figure 17.24 Steady-state heat flow due to conduction in a uniform rod.



(a) Heat current H

(b) Doubling the cross-sectional area of the conductor doubles the heat current (H is proportional to A).

(c) Doubling the length of the conductor halves the heat current (H is inversely proportional to L).

TABLE 17.5 Thermal Conductivities

Substance	k (W/m · K)
Metals	
Aluminum	205.0
Brass	109.0
Copper	385.0
Lead	34.7
Mercury	8.3
Silver	406.0
Steel	50.2
Solids (representative values)	
Brick, insulating	0.15
Brick, red	0.6
Concrete	0.8
Cork	0.04
Felt	0.04
Fiberglass	0.04
Glass	0.8
Ice	1.6
Rock wool	0.04
Styrofoam	0.027
Wood	0.12–0.04
Gases	
Air	0.024
Argon	0.016
Helium	0.14
Hydrogen	0.14
Oxygen	0.023

BIO APPLICATION Fur Versus Blubber The fur of an arctic fox is a good thermal insulator because it traps air, which has a low thermal conductivity k. (The value $k = 0.04$ W/m·K for fur is higher than for air, $k = 0.024$ W/m·K, because fur also includes solid hairs.) The layer of fat beneath a bowhead whale's skin, called blubber, has six times the thermal conductivity of fur ($k = 0.24$ W/m·K). So a 6 cm thickness of blubber ($L = 6$ cm) is required to give the same insulation as 1 cm of fur.

The thermal conductivity of "dead" (nonmoving) air is very small. A wool sweater keeps you warm because it traps air between the fibers. Many insulating materials such as Styrofoam and fiberglass are mostly dead air.

If the temperature varies in a nonuniform way along the length of the conducting rod, we introduce a coordinate x along the length and generalize the temperature gradient to be dT/dx. The corresponding generalization of Eq. (17.21) is

$$H = \frac{dQ}{dt} = -kA\frac{dT}{dx} \qquad (17.22)$$

The negative sign indicates that heat flows in the direction of *decreasing* temperature. If temperature increases with increasing x, then $dT/dx > 0$ and $H < 0$; the negative value of H in this case means that heat flows in the negative x-direction, from high to low temperature.

For thermal insulation in buildings, engineers use the concept of **thermal resistance,** denoted by R. The thermal resistance R of a slab of material with area A is defined so that the heat current H through the slab is

$$H = \frac{A(T_H - T_C)}{R} \qquad (17.23)$$

where T_H and T_C are the temperatures on the two sides of the slab. Comparing this with Eq. (17.21), we see that R is given by

$$R = \frac{L}{k} \qquad (17.24)$$

where L is the thickness of the slab. The SI unit of R is $1\ \text{m}^2 \cdot \text{K/W}$. In the units used for commercial insulating materials in the United States, H is expressed in Btu/h, A is in ft^2, and $T_H - T_C$ in F°. $(1\ \text{Btu/h} = 0.293\ \text{W.})$ The units of R are then $\text{ft}^2 \cdot \text{F}° \cdot \text{h/Btu}$, though values of R are usually quoted without units; a 6-inch-thick layer of fiberglass has an R value of 19 (that is, $R = 19\ \text{ft}^2 \cdot \text{F}° \cdot \text{h/Btu}$), a 2-inch-thick slab of polyurethane foam has an R value of 12, and so on. Doubling the thickness doubles the R value. Common practice in new construction in severe northern climates is to specify R values of around 30 for exterior walls and ceilings. When the insulating material is in layers, such as a plastered wall, fiberglass insulation, and wood exterior siding, the R values are additive. Do you see why?

PROBLEM-SOLVING STRATEGY 17.3 Heat Conduction

IDENTIFY *the relevant concepts:* Heat conduction occurs whenever two objects at different temperatures are placed in contact.

SET UP *the problem* using the following steps:

1. Identify the direction of heat flow (from hot to cold). In Eq. (17.21), L is measured along this direction, and A is an area perpendicular to this direction. You can often approximate an irregular-shaped container with uniform wall thickness as a flat slab with the same thickness and total wall area.
2. List the known and unknown quantities and identify the target variable.

EXECUTE *the solution* as follows:

1. If heat flows through a single object, use Eq. (17.21) to solve for the target variable.
2. If the heat flows through two different materials in succession (in *series*), the temperature T at the interface between them is intermediate between T_H and T_C, so that the temperature differences across the two materials are $(T_H - T)$ and $(T - T_C)$. In steady-state heat flow, the same heat must pass through both materials, so the heat current H must be the *same* in both materials.
3. If heat flows through two or more *parallel* paths, then the total heat current H is the sum of the currents $H_1, H_2, \ldots$ for the separate paths. An example is heat flow from inside a room to outside, both through the glass in a window and through the surrounding wall. In parallel heat flow the temperature difference is the same for each path, but L, A, and k may be different for each path.
4. Be consistent with units. If k is expressed in W/m·K, for example, use distances in meters, heat in joules, and T in kelvins.

EVALUATE *your answer:* Are the results physically reasonable?

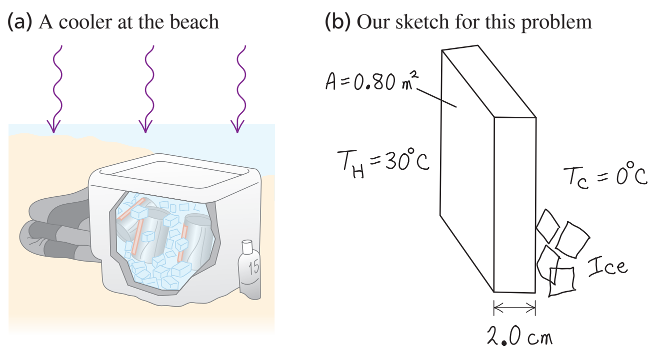
EXAMPLE 17.11 Conduction into a picnic cooler

WITH VARIATION PROBLEMS

A Styrofoam cooler (**Fig. 17.25a**) has total wall area (including the lid) of 0.80 m^2 and wall thickness 2.0 cm. It is filled with ice, water, and cans of Omni-Cola, all at 0°C. What is the rate of heat flow into the cooler if the temperature of the outside wall is 30°C? How much ice melts in 3 hours?

IDENTIFY and SET UP The target variables are the heat current H and the mass m of ice melted. We use Eq. (17.21) to determine H and Eq. (17.20) to determine m.

Figure **17.25** Conduction of heat across the walls of a Styrofoam cooler.



EXECUTE We assume that the total heat flow is the same as it would be through a flat Styrofoam slab of area 0.80 m^2 and thickness 2.0 cm = 0.020 m (Fig. 17.25b). We find k from Table 17.5. From Eq. (17.21),

$$H = kA\frac{T_H - T_C}{L} = (0.027\ \mathrm{W/m\cdot K})(0.80\ \mathrm{m}^2)\frac{30°\mathrm{C} - 0°\mathrm{C}}{0.020\ \mathrm{m}}$$
$$= 32.4\ \mathrm{W} = 32.4\ \mathrm{J/s}$$

The total heat flow is $Q = Ht$, with $t = 3\ \mathrm{h} = 10{,}800\ \mathrm{s}$. From Table 17.4, the heat of fusion of ice is $L_f = 3.34 \times 10^5\ \mathrm{J/kg}$, so from Eq. (17.20) the mass of ice that melts is

$$m = \frac{Q}{L_f} = \frac{(32.4\ \mathrm{J/s})(10{,}800\ \mathrm{s})}{3.34 \times 10^5\ \mathrm{J/kg}} = 1.0\ \mathrm{kg}$$

EVALUATE The low heat current is a result of the low thermal conductivity of Styrofoam.

KEYCONCEPT If a temperature difference is maintained between the two sides of an object of thickness L and cross-sectional area A, there will be a steady heat current due to conduction from the high-temperature side to the low-temperature side. This conduction heat current is proportional to the temperature difference and to the ratio A/L.

EXAMPLE 17.12 Conduction through two bars I

WITH VARIATION PROBLEMS

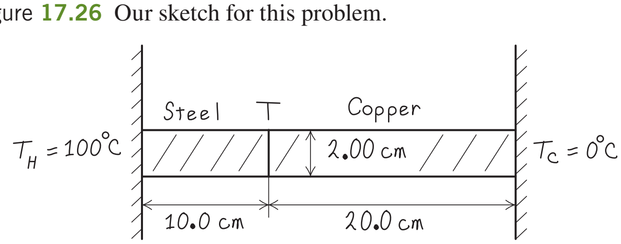
A steel bar 10.0 cm long is welded end to end to a copper bar 20.0 cm long. Each bar has a square cross section, 2.00 cm on a side. The free end of the steel bar is kept at 100°C by placing it in contact with steam, and the free end of the copper bar is kept at 0°C by placing it in contact with ice. Both bars are perfectly insulated on their sides. Find the steady-state temperature at the junction of the two bars and the total rate of heat flow through the bars.

IDENTIFY and SET UP **Figure 17.26** shows the situation. The heat currents in these end-to-end bars must be the same (see Problem-Solving Strategy 17.3). We are given "hot" and "cold" temperatures $T_H = 100°\mathrm{C}$ and $T_C = 0°\mathrm{C}$. With subscripts S for steel and Cu for copper, we write Eq. (17.21) separately for the heat currents H_S and H_{Cu} and set the resulting expressions equal to each other.

EXECUTE Setting $H_S = H_{Cu}$, we have from Eq. (17.21)

$$H_S = k_S A\frac{T_H - T}{L_S} = H_{Cu} = k_{Cu} A\frac{T - T_C}{L_{Cu}}$$

Figure **17.26** Our sketch for this problem.



We divide out the equal cross-sectional areas A and solve for T:

$$T = \frac{\dfrac{k_S}{L_S}T_H + \dfrac{k_{Cu}}{L_{Cu}}T_C}{\left(\dfrac{k_S}{L_S} + \dfrac{k_{Cu}}{L_{Cu}}\right)}$$

Substituting $L_S = 10.0\ \mathrm{cm}$ and $L_{Cu} = 20.0\ \mathrm{cm}$, the given values of T_H and T_C, and the values of k_S and k_{Cu} from Table 17.5, we find $T = 20.7°\mathrm{C}$.

We can find the total heat current by substituting this value of T into either the expression for H_S or the one for H_{Cu}:

$$H_S = (50.2\ \mathrm{W/m\cdot K})(0.0200\ \mathrm{m})^2\frac{100°\mathrm{C} - 20.7°\mathrm{C}}{0.100\ \mathrm{m}}$$
$$= 15.9\ \mathrm{W}$$

$$H_{Cu} = (385\ \mathrm{W/m\cdot K})(0.0200\ \mathrm{m})^2\frac{20.7°\mathrm{C}}{0.200\ \mathrm{m}} = 15.9\ \mathrm{W}$$

EVALUATE Even though the steel bar is shorter, the temperature drop across it is much greater (from 100°C to 20.7°C) than across the copper bar (from 20.7°C to 0°C). That's because steel is a much poorer conductor than copper.

KEYCONCEPT When there is a steady heat flow by conduction through two materials in succession, the heat current is the same in both materials: No energy is lost in going from one material to the next.

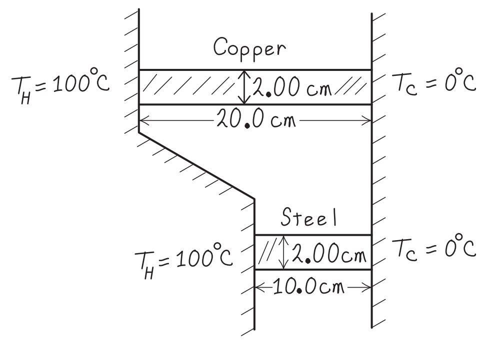
EXAMPLE 17.13 Conduction through two bars II

WITH VARIATION PROBLEMS

Suppose the two bars of Example 17.12 are separated. One end of each bar is kept at 100°C and the other end of each bar is kept at 0°C. What is the *total* heat current in the two bars?

IDENTIFY and SET UP **Figure 17.27** shows the situation. For each bar, $T_H - T_C = 100°C - 0°C = 100$ K. The total heat current is the sum of the currents in the two bars, $H_S + H_{Cu}$.

Figure 17.27 Our sketch for this problem.



EXECUTE We write the heat currents for the two rods individually, and then add them to get the total heat current:

$$H = H_S + H_{Cu} = k_S A \frac{T_H - T_C}{L_S} + k_{Cu} A \frac{T_H - T_C}{L_{Cu}}$$

$$= (50.2 \text{ W/m}\cdot\text{K})(0.0200 \text{ m})^2 \frac{100 \text{ K}}{0.100 \text{ m}}$$

$$+ (385 \text{ W/m}\cdot\text{K})(0.0200 \text{ m})^2 \frac{100 \text{ K}}{0.200 \text{ m}}$$

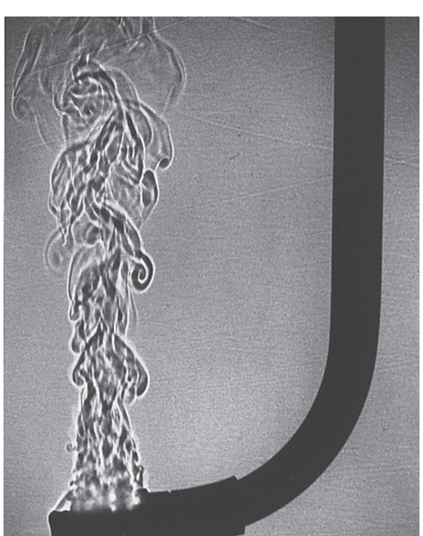
$$= 20.1 \text{ W} + 77.0 \text{ W} = 97.1 \text{ W}$$

EVALUATE The heat flow in the copper bar is much greater than that in the steel bar, even though it is longer, because the thermal conductivity of copper is much larger. The total heat flow is greater than in Example 17.12 because the total cross section for heat flow is greater and because the full 100 K temperature difference appears across each bar.

KEYCONCEPT Even if two different objects have the same constant temperature difference between their ends, the heat current H through the two objects can be different. The value of H depends on the dimensions of the object and the thermal conductivity of the material of the object.

Convection

Figure 17.28 A heating element in the tip of this submerged tube warms the surrounding water, producing a complex pattern of free convection.



Convection is the transfer of heat by mass motion of a fluid from one region of space to another. Familiar examples include hot-air and hot-water home heating systems, the cooling system of an automobile engine, and the flow of blood in the body. If the fluid is circulated by a blower or pump, the process is called *forced convection;* if the flow is caused by differences in density due to thermal expansion, such as hot air rising, the process is called *free convection* (**Fig. 17.28**).

Free convection in the atmosphere plays a dominant role in determining the daily weather, and convection in the oceans is an important global heat-transfer mechanism. On a smaller scale, soaring hawks and glider pilots make use of thermal updrafts from the warm earth. The most important mechanism for heat transfer within the human body (needed to maintain nearly constant temperature in various environments) is *forced* convection of blood, with the heart as the pump.

Convective heat transfer is a very complex process, and there is no simple equation to describe it. Here are a few experimental facts:

1. The heat current due to convection is directly proportional to the surface area. That's why radiators and cooling fins, which use convection to transfer heat, have large surface areas.

2. The viscosity of fluids slows natural convection near a stationary surface. For air, this gives rise to a surface film that on a vertical surface typically has about the same insulating value as 1.3 cm of plywood (R value $= 0.7$). Forced convection decreases the thickness of this film, increasing the rate of heat transfer. This is the reason for the "wind-chill factor"; you get cold faster in a cold wind than in still air with the same temperature.

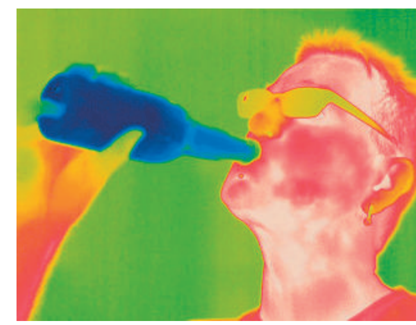
3. The heat current due to free convection is found to be approximately proportional to the $\frac{5}{4}$ power of the temperature difference between the surface and the main body of fluid.

Radiation

Radiation is the transfer of heat by electromagnetic waves such as visible light, infrared, and ultraviolet radiation. Everyone has felt the warmth of the sun's radiation and the intense heat from a charcoal grill or the glowing coals in a fireplace. Most of the heat from these very hot objects reaches you not by conduction or convection in the intervening air but by *radiation*. This heat transfer would occur even if there were nothing but vacuum between you and the source of heat.

Every object, even at ordinary temperatures, emits energy in the form of electromagnetic radiation. Around 20°C, nearly all the energy is carried by infrared waves with wavelengths much longer than those of visible light (see Fig. 17.4 and **Fig. 17.29**). As the temperature rises, the wavelengths shift to shorter values. At 800°C, an object emits enough visible radiation to appear "red-hot," although even at this temperature most of the energy is carried by infrared waves. At 3000°C, the temperature of an incandescent lamp filament, the radiation contains enough visible light that the object appears "white-hot."

Figure **17.29** This false-color infrared photograph reveals radiation emitted by various parts of the man's body. The strongest emission (colored red) comes from the warmest areas, while there is very little emission from the bottle of cold beverage.



The rate of energy radiation from a surface is proportional to the surface area A and to the fourth power of the absolute (Kelvin) temperature T. The rate also depends on the nature of the surface; we describe this dependence by a quantity e called **emissivity.** A dimensionless number between 0 and 1, e is the ratio of the rate of radiation from a particular surface to the rate of radiation from an equal area of an ideal radiating surface at the same temperature. Emissivity also depends somewhat on temperature. Thus we can to nine significant figures, its express the heat current $H = dQ/dt$ due to radiation from a surface as

Area of emitting surface

Heat current in radiation Emissivity of surface

$$H = Ae\sigma T^4 \tag{17.25}$$

Stefan–Boltzmann constant Absolute temperature of surface

This relationship is called the **Stefan–Boltzmann law** in honor of its late-19th-century discoverers. The **Stefan–Boltzmann constant** σ (Greek sigma) is a fundamental constant; to nine significant figures, its numerical value is

$$\sigma = 5.67037442 \times 10^{-8}\ \text{W/m}^2 \cdot \text{K}^4$$

You should check unit consistency in Eq. (17.25). Emissivity (e) is often larger for dark surfaces than for light ones. The emissivity of a smooth copper surface is about 0.3, but e for a dull black surface can be close to unity.

EXAMPLE 17.14 Heat transfer by radiation

WITH VARIATION PROBLEMS

A thin, square steel plate, 10 cm on a side, is heated in a blacksmith's forge to 800°C. If the emissivity is 0.60, what is the total rate of radiation of energy from the plate?

IDENTIFY and SET UP The target variable is H, the rate of emission of energy from the plate's two surfaces. We use Eq. (17.25) to calculate H.

EXECUTE The total surface area is $2(0.10\ \text{m})^2 = 0.020\ \text{m}^2$, and $T = 800°\text{C} = 1073\ \text{K}$. Then Eq. (17.25) gives

$$\begin{aligned} H &= Ae\sigma T^4 \\ &= (0.020\ \text{m}^2)(0.60)(5.67 \times 10^{-8}\ \text{W/m}^2 \cdot \text{K}^4)(1073\ \text{K})^4 = 900\ \text{W} \end{aligned}$$

EVALUATE The nearby blacksmith will easily feel the heat radiated from this plate.

KEYCONCEPT All objects emit energy in the form of electromagnetic radiation due to their temperature. The heat current of this radiation is proportional to the object's surface area, to the emissivity of its surface, and to the fourth power of the object's Kelvin temperature.

Radiation and Absorption

While an object at absolute temperature T is radiating, its surroundings at temperature T_s are also radiating, and the object *absorbs* some of this radiation. If it is in thermal equilibrium with its surroundings, $T = T_s$ and the rates of radiation and absorption must be equal. For this to be true, the rate of absorption must be given in general by $H = Ae\sigma T_s^{\,4}$.

Then the *net* rate of radiation from an object at temperature T with surroundings at temperature T_s is $Ae\sigma T^4 - Ae\sigma T_s^4$, or

$$\underbrace{H_{net}}_{\text{Net heat current in radiation}} = \underbrace{A}_{\text{Area of emitting surface}}\underbrace{e}_{\text{Emissivity of surface}}\underbrace{\sigma}_{\text{Stefan–Boltzmann constant}}(\underbrace{T^4 - T_s^4}_{\text{Absolute temperatures of surface } (T) \text{ and surroundings } (T_s)}) \quad (17.26)$$

In Eq. (17.26) a positive value of H means a net heat flow *out of* the object. This will be the case if $T > T_s$.

EXAMPLE 17.15 Radiation from the human body

WITH VARIATION PROBLEMS

What is the total rate of radiation of energy from a human body with surface area 1.20 m^2 and surface temperature 30°C = 303 K? If the surroundings are at a temperature of 20°C, what is the *net* rate of radiative heat loss from the body? The emissivity of the human body is very close to unity, irrespective of skin pigmentation.

IDENTIFY and SET UP We must consider both the radiation that the body emits and the radiation that it absorbs from its surroundings. Equation (17.25) gives the rate of radiation of energy from the body, and Eq. (17.26) gives the net rate of heat loss.

EXECUTE Taking $e = 1$ in Eq. (17.25), we find that the body radiates at a rate

$$\begin{aligned} H &= Ae\sigma T^4 \\ &= (1.20\ \text{m}^2)(1)(5.67 \times 10^{-8}\ \text{W/m}^2 \cdot \text{K}^4)(303\ \text{K})^4 = 574\ \text{W} \end{aligned}$$

This loss is partly offset by absorption of radiation, which depends on the temperature of the surroundings. From Eq. (17.26), the *net* rate of radiative energy transfer is

$$\begin{aligned} H_{net} &= Ae\sigma(T^4 - T_s^4) \\ &= (1.20\ \text{m}^2)(1)(5.67 \times 10^{-8}\ \text{W/m}^2 \cdot \text{K}^4)[(303\ \text{K})^4 - (293\ \text{K})^4] \\ &= 72\ \text{W} \end{aligned}$$

EVALUATE The value of H_{net} is positive because the body is losing heat to its colder surroundings.

KEYCONCEPT An object at Kelvin temperature T emits electromagnetic radiation but also absorbs radiation from its surroundings at Kelvin temperature T_s. The *net* heat current is proportional to the object's surface area, to the emissivity of its surface, and to the difference between T^4 and T_s^4.

Applications of Radiation

Heat transfer by radiation is important in some surprising places. A premature baby in an incubator can be cooled dangerously by radiation if the walls of the incubator happen to be cold, even when the *air* in the incubator is warm. Some incubators regulate the air temperature by measuring the baby's skin temperature.

An object that is a good absorber must also be a good emitter. An ideal radiator, with emissivity $e = 1$, is also an ideal absorber, absorbing *all* of the radiation that strikes it. Such an ideal surface is called an ideal black body or simply a **blackbody.** Conversely, an ideal *reflector,* which absorbs *no* radiation at all, is also a very ineffective radiator.

This is the reason for the silver coatings on vacuum ("Thermos") bottles, invented by Sir James Dewar (1842–1923). A vacuum bottle has double glass walls. The air is pumped out of the spaces between the walls; this eliminates nearly all heat transfer by conduction and convection. The silver coating on the walls reflects most of the radiation from the contents back into the container, and the wall itself is a very poor emitter. Thus a vacuum bottle can keep coffee or soup hot for several hours. The Dewar flask, used to store very cold liquefied gases, is exactly the same in principle.

Radiation, Climate, and Climate Change

Our planet constantly absorbs radiation coming from the sun. In thermal equilibrium, the rate at which our planet absorbs solar radiation must equal the rate at which it emits radiation into space. The presence of an atmosphere on our planet has a significant effect on this equilibrium.

Most of the radiation emitted by the sun (which has a surface temperature of 5800 K) is in the visible part of the spectrum, to which our atmosphere is transparent. But the average surface temperature of the earth is only 287 K (14°C). Hence most of the radiation that our planet emits into space is infrared radiation, just like the radiation from the person shown in

Fig. 17.29. However, our atmosphere is *not* completely transparent to infrared radiation. This is because our atmosphere contains carbon dioxide (CO_2), which is its fourth most abundant constituent (after nitrogen, oxygen, and argon). Molecules of CO_2 in the atmosphere *absorb* some of the infrared radiation coming upward from the surface. They then re-radiate the absorbed energy, but some of the re-radiated energy is directed back down toward the surface instead of escaping into space. In order to maintain thermal equilibrium, the earth's surface must compensate for this by increasing its temperature T and hence its total rate of radiating energy (which is proportional to T^4). This phenomenon, called the **greenhouse effect,** makes our planet's surface temperature about 33°C higher than it would be if there were no atmospheric CO_2. If CO_2 were absent, the earth's average surface temperature would be below the freezing point of water, and life as we know it would be impossible.

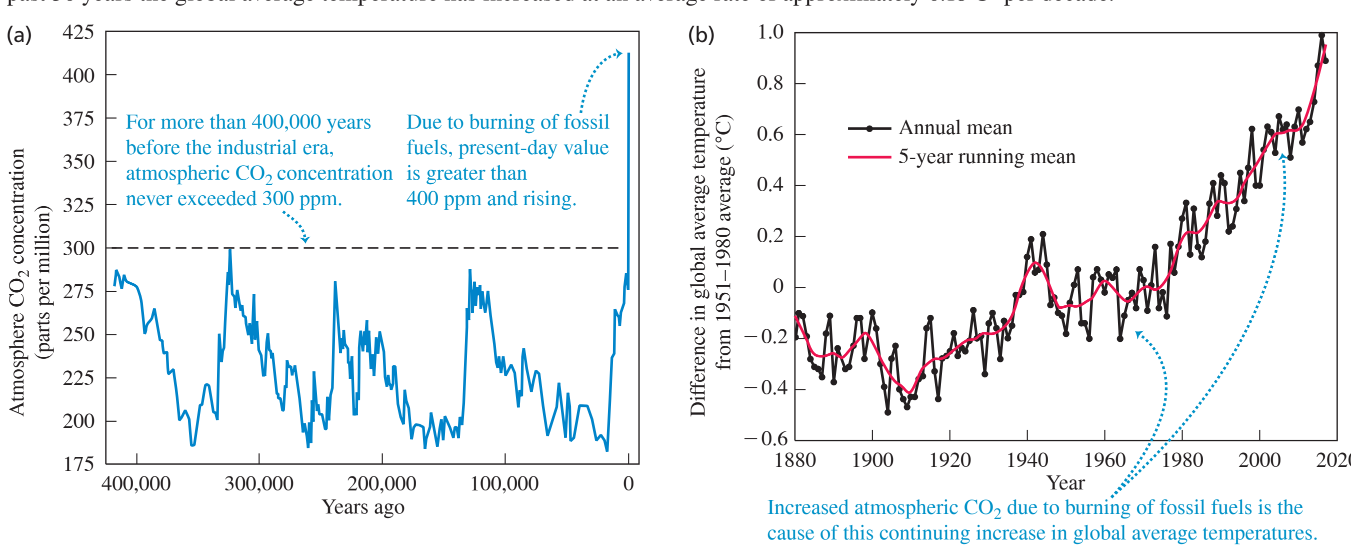
While atmospheric CO_2 has benefits, too much of it can have extremely negative consequences. Measurements of air trapped in ancient Antarctic ice show that over the past 650,000 years CO_2 has constituted less than 300 parts per million of our atmosphere. Since the beginning of the industrial age, however, the burning of fossil fuels such as coal and petroleum has elevated the atmospheric CO_2 concentration to unprecedented levels (**Fig. 17.30a**). As a consequence, since the 1950s the global average surface temperature has increased by 0.9°C and the earth has experienced the hottest years ever recorded (Fig. 17.30b). If we continue to consume fossil fuels at the same rate, by 2050 the atmospheric CO_2 concentration will reach 600 parts per million, well off the scale of Fig. 17.30a. The resulting temperature increase will have dramatic effects on global climate. In polar regions massive quantities of ice will melt and run from solid land to the sea, thus raising ocean levels worldwide and threatening the homes and lives of hundreds of millions of people who live near the coast. Coping with these threats is one of the greatest challenges facing 21st-century civilization.

TEST YOUR UNDERSTANDING OF SECTION 17.7 A room has one wall made of concrete, one wall made of copper, and one wall made of steel. All of the walls are the same size and at the same temperature of 20°C. Which wall feels coldest to the touch? (i) The concrete wall; (ii) the copper wall; (iii) the steel wall; (iv) all three walls feel equally cold.

ANSWER

(ii) When you touch one of the walls, heat flows from your hand to the lower-temperature wall. The more rapidly heat flows from your hand, the colder you'll feel. Equation (17.21) shows that the rate of heat flow is proportional to the thermal conductivity k. From Table 17.5, copper has a much higher thermal conductivity ($385.0\ \mathrm{W/m \cdot K}$) than steel ($50.2\ \mathrm{W/m \cdot K}$) or concrete ($0.8\ \mathrm{W/m \cdot K}$), and so the copper wall feels the coldest.

Figure **17.30** (a) Due to humans burning fossil fuels, the concentration of carbon dioxide in the atmosphere is now more than 33% greater than in the pre-industrial era. (b) Due to the increased CO_2 concentration, during the past 50 years the global average temperature has increased at an average rate of approximately 0.18 C° per decade.



CHAPTER 17 SUMMARY

Temperature and temperature scales: Two objects in thermal equilibrium must have the same temperature. A conducting material between two objects permits them to interact and come to thermal equilibrium; an insulating material impedes this interaction.

The Celsius and Fahrenheit temperature scales are based on the freezing ($0°C = 32°F$) and boiling ($100°C = 212°F$) temperatures of water. One Celsius degree equals $\frac{9}{5}$ Fahrenheit degrees. (See Example 17.1.)

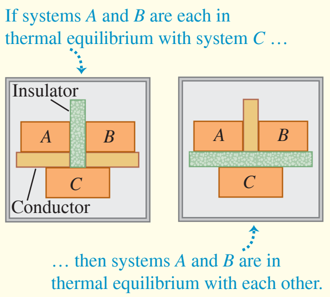
The Kelvin scale has its zero at the extrapolated zero-pressure temperature for a gas thermometer, $-273.15°C = 0$ K. In the gas-thermometer scale, the ratio of two temperatures T_1 and T_2 is defined to be equal to the ratio of the two corresponding gas-thermometer pressures p_1 and p_2.

$$T_F = \tfrac{9}{5}T_C + 32° \quad (17.1)$$

$$T_C = \tfrac{5}{9}(T_F - 32°) \quad (17.2)$$

$$T_K = T_C + 273.15 \quad (17.3)$$

$$\frac{T_2}{T_1} = \frac{p_2}{p_1} \quad (17.4)$$



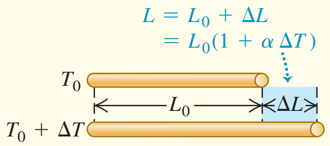
Thermal expansion and thermal stress: A temperature change ΔT causes a change in any linear dimension L_0 of a solid object. The change ΔL is approximately proportional to L_0 and ΔT. Similarly, a temperature change causes a change ΔV in the volume V_0 of any solid or liquid; ΔV is approximately proportional to V_0 and ΔT. The quantities α and β are the coefficients of linear expansion and volume expansion, respectively. For solids, $\beta = 3\alpha$. (See Examples 17.2 and 17.3.)

When a material is cooled or heated and held so it cannot contract or expand, it is under a tensile stress F/A. (See Example 17.4.)

$$\Delta L = \alpha L_0 \,\Delta T \quad (17.6)$$

$$\Delta V = \beta V_0 \,\Delta T \quad (17.8)$$

$$\frac{F}{A} = -Y\alpha\,\Delta T \quad (17.12)$$



Heat, phase changes, and calorimetry: Heat is energy in transit from one object to another as a result of a temperature difference. Equations (17.13) and (17.18) give the quantity of heat Q required to cause a temperature change ΔT in a quantity of material with mass m and specific heat c (alternatively, with number of moles n and molar heat capacity $C = Mc$, where M is the molar mass and $m = nM$). When heat is added to an object, Q is positive; when it is removed, Q is negative. (See Examples 17.5 and 17.6.)

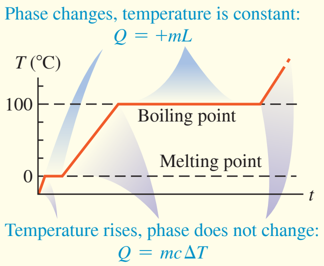
To change a mass m of a material to a different phase at the same temperature (such as liquid to vapor), a quantity of heat given by Eq. (17.20) must be added or subtracted. Here L is the heat of fusion, vaporization, or sublimation.

In an isolated system whose parts interact by heat exchange, the algebraic sum of the Q's for all parts of the system must be zero. (See Examples 17.7–17.10.)

$$Q = mc\,\Delta T \quad (17.13)$$

$$Q = nC\,\Delta T \quad (17.18)$$

$$Q = \pm mL \quad (17.20)$$



Conduction, convection, and radiation: Conduction is the transfer of heat within materials without bulk motion of the materials. The heat current H depends on the area A through which the heat flows, the length L of the heat-flow path, the temperature difference $(T_H - T_C)$, and the thermal conductivity k of the material. (See Examples 17.11–17.13.)

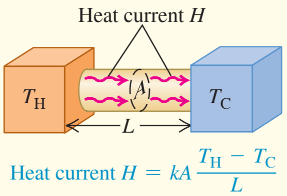
Convection is a complex heat-transfer process that involves mass motion from one region to another.

Radiation is energy transfer through electromagnetic radiation. The radiation heat current H depends on the surface area A, the emissivity e of the surface (a pure number between 0 and 1), and the Kelvin temperature T. Here σ is the Stefan–Boltzmann constant. The *net* radiation heat current H_{net} from an object at temperature T to its surroundings at temperature T_s depends on both T and T_s. (See Examples 17.14 and 17.15.)

$$H = \frac{dQ}{dt} = kA\frac{T_H - T_C}{L} \quad (17.21)$$

$$H = Ae\sigma T^4 \quad (17.25)$$

$$H_{net} = Ae\sigma(T^4 - T_s^4) \quad (17.26)$$



Chapter 17 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 17.2, 17.3, and 17.4 (Section 17.4) before attempting these problems.

VP17.4.1 A metal rod is 0.500 m in length at a temperature of 15.0°C. When you raise its temperature to 37.0°C, its length increases by 0.220 mm. (a) What is the coefficient of linear expansion of the metal? (b) If a second rod of the same metal has length 0.300 m at 25.0°C, how will its length change if the temperature drops to −20.0°C?

VP17.4.2 A copper mug that can hold 250 cm^3 of liquid is filled to the brim with ethanol at 20.0°C. If you lower the temperature of the mug and ethanol to −50.0°C, what is the maximum additional volume of ethanol you can add to the mug without spilling any? (See Table 17.2. Ethanol remains a liquid at temperatures down to −114°C.)

VP17.4.3 A cylindrical brass rod is 10.0 cm in length and 0.500 cm in radius at 25.0°C. How much force do you have to apply to each end of the rod to maintain its length when the temperature is decreased to 13.0°C? Are the required forces tensile or compressive? (See Table 17.1. Brass has Young's modulus 9.0×10^{10} Pa.)

VP17.4.4 A rod made of metal A is attached end to end to another rod made of metal B, making a combined rod of overall length L. The coefficients of linear expansion of metals A and B are α_A and α_B, respectively, and $\alpha_B > \alpha_A$. When the temperature of the combined rod is increased by ΔT, the overall length increases by ΔL. What was the initial length of the rod of metal A?

Be sure to review EXAMPLES 17.7, 17.8, and 17.9 (Section 17.6) before attempting these problems.

VP17.9.1 You place a piece of aluminum at 250.0°C in 5.00 kg of liquid water at 20.0°C. None of the water boils, and the final temperature of the water and aluminum is 22.0°C. What is the mass of the piece of aluminum? Assume no heat is exchanged with the container that holds the water. (See Table 17.3.)

VP17.9.2 You place an ice cube of mass 7.50×10^{-3} kg and temperature 0.00°C on top of a copper cube of mass 0.460 kg. All of the ice melts, and the final equilibrium temperature of the two substances is 0.00°C. What was the initial temperature of the copper cube? Assume no heat is exchanged with the surroundings. (See Tables 17.3 and 17.4.)

VP17.9.3 You have 1.60 kg of liquid ethanol at 28.0°C that you wish to cool. What mass of ice at initial temperature −5.00°C should you add to the ethanol so that all of the ice melts and the resulting ethanol–water mixture has temperature 10.0°C? Assume no heat is exchanged with the container that holds the ethanol. (See Tables 17.3 and 17.4.)

VP17.19.4 You put a silver ingot of mass 1.25 kg and initial temperature 315°C in contact with 0.250 kg of ice at initial temperature −8.00°C. Assume no heat is exchanged with the surroundings. (a) What is the final equilibrium temperature? (b) What fraction of the ice melts? (See Tables 17.3 and 17.4.)

Be sure to review EXAMPLES 17.11, 17.12, 17.13, 17.14, and 17.15 (Section 17.7) before attempting these problems.

VP17.15.1 A square pane of glass 0.500 m on a side is 6.00 mm thick. When the temperatures on the two sides of the glass are 25.0°C and −10.0°C, the heat current due to conduction through the glass is 1.10×10^3 W. (a) What is the thermal conductivity of the glass? (b) If the thickness of the glass is increased to 9.00 mm, what will be the heat current?

VP17.15.2 A brass rod and a lead rod, each 0.250 m long and each with cross-sectional area 2.00×10^{-4} m^2, are joined end to end to make a composite rod of overall length 0.500 m. The free end of the brass rod is maintained at a high temperature, and the free end of the lead rod is maintained at a low temperature. The temperature at the junction of the two rods is 185°C, and the heat current due to conduction through the composite rod is 6.00 W. What are the temperatures of (a) the free end of the brass rod and (b) the free end of the lead rod? (See Table 17.5.)

VP17.15.3 The emissivity of the surface of a star is approximately 1. The star Sirius A emits electromagnetic radiation at a rate of 9.7×10^{27} W and has a surface temperature of 9940 K. What is the radius of Sirius in meters and as a multiple of the sun's radius (6.96×10^8 m)?

VP17.15.4 A building in the desert is made of concrete blocks (emissivity 0.91) and has an exposed surface area of 525 m^2. If the building is maintained at 20.0°C but the temperature on a hot desert night is 35.0°C, what is the net rate at which the building absorbs energy by radiation?

BRIDGING PROBLEM Steady-State Heat Flow: Radiation and Conduction

One end of a solid cylindrical copper rod 0.200 m long and 0.0250 m in radius is inserted into a large block of solid hydrogen at its melting temperature, 13.84 K. The other end is blackened and exposed to thermal radiation from surrounding walls at 500.0 K. (Some telescopes in space employ a similar setup. A solid refrigerant keeps the telescope very cold—required for proper operation—even though it is exposed to direct sunlight.) The sides of the rod are insulated, so no energy is lost or gained except at the ends of the rod. (a) When equilibrium is reached, what is the temperature of the blackened end? The thermal conductivity of copper at temperatures near 20 K is 1670 W/m · K. (b) At what rate (in kg/h) does the solid hydrogen melt?

SOLUTION GUIDE

IDENTIFY and SET UP

1. Draw a sketch of the situation, showing all relevant dimensions.
2. List the known and unknown quantities, and identify the target variables.
3. In order for the rod to be in equilibrium, how must the radiation heat current from the walls into the blackened end of the rod compare to the conduction heat current from this end to the other end and into the solid hydrogen? Use your answers to select the appropriate equations for part (a).
4. How does the heat current from the rod into the hydrogen determine the rate at which the hydrogen melts? (*Hint:* See Table 17.4.) Use your answer to select the appropriate equations for part (b).

EXECUTE

5. Solve for the temperature of the blackened end of the rod. (*Hint:* Since copper is an excellent conductor of heat at low temperature, you can assume that the temperature of the blackened end is only slightly higher than 13.84 K.)
6. Use your result from step 5 to find the rate at which the hydrogen melts.

EVALUATE

7. Is your result from step 5 consistent with the hint in that step?
8. How would your results from steps 5 and 6 be affected if the rod had twice the radius?

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q17.1 Explain why it would not make sense to use a full-size glass thermometer to measure the temperature of a thimbleful of hot water.

Q17.2 If you heat the air inside a rigid, sealed container until its Kelvin temperature doubles, the air pressure in the container will also double. Is the same thing true if you double the Celsius temperature of the air in the container? Explain.

Q17.3 Many automobile engines have cast-iron cylinders and aluminum pistons. What kinds of problems could occur if the engine gets too hot? (The coefficient of volume expansion of cast iron is approximately the same as that of steel.)

Q17.4 Why do frozen water pipes burst? Would a mercury thermometer break if the temperature went below the freezing temperature of mercury? Why or why not?

Q17.5 Two objects made of the same material have the same external dimensions and appearance, but one is solid and the other is hollow. When their temperature is increased, is the overall volume expansion the same or different? Why?

Q17.6 Why is it sometimes possible to loosen caps on screw-top bottles by dipping the capped bottle briefly into hot water?

Q17.7 The inside of an oven is at a temperature of 200°C (392°F). You can put your hand in the oven without injury as long as you don't touch anything. But since the air inside the oven is also at 200°C, why isn't your hand burned just the same?

Q17.8 A newspaper article about the weather states that "the temperature of an object measures how much heat the object contains." Is this description correct? Why or why not?

Q17.9 A student asserts that a suitable unit for specific heat is $1\ \mathrm{m^2/s^2 \cdot C^\circ}$. Is she correct? Why or why not?

Q17.10 In some household air conditioners used in dry climates, air is cooled by blowing it through a water-soaked filter, evaporating some of the water. How does this cool the air? Would such a system work well in a high-humidity climate? Why or why not?

Q17.11 The units of specific heat c are $\mathrm{J/kg \cdot K}$, but the units of heat of fusion L_f or heat of vaporization L_v are simply J/kg. Why do the units of L_f and L_v not include a factor of $(\mathrm{K})^{-1}$ to account for a temperature change?

Q17.12 Why is a hot, humid day in the tropics generally more uncomfortable for human beings than a hot, dry day in the desert?

Q17.13 A piece of aluminum foil used to wrap a potato for baking in a hot oven can usually be handled safely within a few seconds after the potato is removed from the oven. The same is not true of the potato, however! Give two reasons for this difference.

Q17.14 Desert travelers sometimes keep water in a canvas bag. Some water seeps through the bag and evaporates. How does this cool the water inside the bag?

Q17.15 When you first step out of the shower, you feel cold. But as soon as you are dry you feel warmer, even though the room temperature does not change. Why?

Q17.16 The climate of regions adjacent to large bodies of water (like the Pacific and Atlantic coasts) usually features a narrower range of temperature than the climate of regions far from large bodies of water (like the prairies). Why?

Q17.17 When water is placed in ice-cube trays in a freezer, why doesn't the water freeze all at once when the temperature has reached 0°C? In fact, the water freezes first in a layer adjacent to the sides of the tray. Why?

Q17.18 Before giving you an injection, a physician swabs your arm with isopropyl alcohol at room temperature. Why does this make your arm feel cold? (*Hint:* The reason is *not* the fear of the injection! The boiling point of isopropyl alcohol is 82.4°C.)

Q17.19 A cold block of metal feels colder than a block of wood at the same temperature. Why? A *hot* block of metal feels hotter than a block of wood at the same temperature. Again, why? Is there any temperature at which the two blocks feel equally hot or cold? What temperature is this?

Q17.20 A person pours a cup of hot coffee, intending to drink it five minutes later. To keep the coffee as hot as possible, should she put cream in it now or wait until just before she drinks it? Explain.

Q17.21 When a freshly baked apple pie has just been removed from the oven, the crust and filling are both at the same temperature. Yet if you sample the pie, the filling will burn your tongue but the crust will not. Why is there a difference? (*Hint:* The filling is moist while the crust is dry.)

Q17.22 Old-time kitchen lore suggests that things cook better (evenly and without burning) in heavy cast-iron pots. What desirable characteristics do such pots have?

Q17.23 In coastal regions in the winter, the temperature over the land is generally colder than the temperature over the nearby ocean; in the summer, the reverse is usually true. Explain. (*Hint:* The specific heat of soil is only 0.2–0.8 times as great as that of water.)

Q17.24 It is well known that a potato bakes faster if a large nail is stuck through it. Why? Does an aluminum nail work better than a steel one? Why or why not? (*Note:* Don't try this in a microwave oven!) There is also a gadget on the market to hasten the roasting of meat; it consists of a hollow metal tube containing a wick and some water. This is claimed to work much better than a solid metal rod. How does it work?

Q17.25 Glider pilots in the Midwest know that thermal updrafts are likely to occur in the vicinity of freshly plowed fields. Why?

Q17.26 Some folks claim that ice cubes freeze faster if the trays are filled with hot water, because hot water cools off faster than cold water. What do you think?

Q17.27 We're lucky that the earth isn't in thermal equilibrium with the sun (which has a surface temperature of 5800 K). But why aren't the two objects in thermal equilibrium?

Q17.28 When energy shortages occur, magazine articles sometimes urge us to keep our homes at a constant temperature day and night to conserve fuel. They argue that when we turn down the heat at night, the walls, ceilings, and other areas cool off and must be reheated in the morning. So if we keep the temperature constant, these parts of the house will not cool off and will not have to be reheated. Does this argument make sense? Would we really save energy by following this advice?

EXERCISES

Section 17.2 Thermometers and Temperature Scales

17.1 • Convert the following Celsius temperatures to Fahrenheit: (a) −62.8°C, the lowest temperature ever recorded in North America (February 3, 1947, Snag, Yukon); (b) 56.7°C, the highest temperature ever recorded in the United States (July 10, 1913, Death Valley, California); (c) 31.1°C, the world's highest average annual temperature (Lugh Ferrandi, Somalia).

17.2 • BIO **Temperatures in Biomedicine.** (a) **Normal body temperature.** The average normal body temperature measured in the mouth is 310 K. What would Celsius and Fahrenheit thermometers read for this temperature? (b) **Elevated body temperature.** During very vigorous exercise, the body's temperature can go as high as 40°C. What would Kelvin and Fahrenheit thermometers read for this temperature? (c) **Temperature difference in the body.** The surface temperature of the body is normally about 7 C° lower than the internal temperature. Express this temperature difference in kelvins and in Fahrenheit degrees. (d) **Blood storage.** Blood stored at 4.0°C lasts safely for about 3 weeks, whereas blood stored at −160°C lasts for 5 years. Express both temperatures on the Fahrenheit and Kelvin scales. (e) **Heat stroke.** If the body's temperature is above 105°F for a prolonged period, heat stroke can result. Express this temperature on the Celsius and Kelvin scales.

17.3 • (a) On January 22, 1943, the temperature in Spearfish, South Dakota, rose from −4.0°F to 45.0°F in just 2 minutes. What was the temperature change in Celsius degrees? (b) The temperature in Browning, Montana, was 44.0°F on January 23, 1916. The next day the temperature plummeted to −56°F. What was the temperature change in Celsius degrees?

Section 17.3 Gas Thermometers and the Kelvin Scale

17.4 • Derive an equation that gives T_K as a function of T_F to the nearest hundredth of a degree. Solve the equation and thereby obtain an equation for T_F as a function of T_K.

17.5 •• You put a bottle of soft drink in a refrigerator and leave it until its temperature has dropped 10.0 K. What is its temperature change in (a) F° and (b) C°?

17.6 • (a) Calculate the one temperature at which Fahrenheit and Celsius thermometers agree with each other. (b) Calculate the one temperature at which Fahrenheit and Kelvin thermometers agree with each other.

17.7 • The pressure of a gas at the triple point of water is 1.35 atm. If its volume remains unchanged, what will its pressure be at the temperature at which CO_2 solidifies?

17.8 • Convert the following Kelvin temperatures to the Celsius and Fahrenheit scales: (a) the midday temperature at the surface of the moon (400 K); (b) the temperature at the tops of the clouds in the atmosphere of Saturn (95 K); (c) the temperature at the center of the sun $(1.55 \times 10^7\ \text{K})$.

17.9 •• **A Constant-Volume Gas Thermometer.** An experimenter using a gas thermometer found the pressure at the triple point of water (0.01°C) to be 4.80×10^4 Pa and the pressure at the normal boiling point (100°C) to be 6.50×10^4 Pa. (a) Assuming that the pressure varies linearly with temperature, use these two data points to find the Celsius temperature at which the gas pressure would be zero (that is, find the Celsius temperature of absolute zero). (b) Does the gas in this thermometer obey Eq. (17.4) precisely? If that equation were precisely obeyed and the pressure at 100°C were 6.50×10^4 Pa, what pressure would the experimenter have measured at 0.01°C? (As we'll learn in Section 18.1, Eq. (17.4) is accurate only for gases at very low density.)

17.10 •• A constant-volume gas thermometer registers an absolute pressure corresponding to 325 mm of mercury when in contact with water at the triple point. What pressure does it read when in contact with water at the normal boiling point?

Section 17.4 Thermal Expansion

17.11 • The Humber Bridge in England has the world's longest single span, 1410 m. Calculate the change in length of the steel deck of the span when the temperature increases from −5.0°C to 18.0°C.

17.12 • One of the tallest buildings in the world is the Taipei 101 in Taiwan, at a height of 1671 feet. Assume that this height was measured on a cool spring day when the temperature was 15.5°C. You could use the building as a sort of giant thermometer on a hot summer day by carefully measuring its height. Suppose you do this and discover that the Taipei 101 is 0.471 foot taller than its official height. What is the temperature, assuming that the building is in thermal equilibrium with the air and that its entire frame is made of steel?

17.13 • A U.S. penny has a diameter of 1.9000 cm at 20.0°C. The coin is made of a metal alloy (mostly zinc) for which the coefficient of linear expansion is $2.6 \times 10^{-5}\ \text{K}^{-1}$. What would its diameter be on a hot day in Death Valley (48.0°C)? On a cold night in the mountains of Greenland (−53°C)?

17.14 • **Ensuring a Tight Fit.** Aluminum rivets used in airplane construction are made slightly larger than the rivet holes and cooled by "dry ice" (solid CO_2) before being driven. If the diameter of a hole is 4.500 mm, what should be the diameter of a rivet at 23.0°C if its diameter is to equal that of the hole when the rivet is cooled to −78.0°C, the temperature of dry ice? Assume that the expansion coefficient remains constant at the value given in Table 17.1.

17.15 •• A copper cylinder is initially at 20.0°C. At what temperature will its volume be 0.150% larger than it is at 20.0°C?

17.16 •• A geodesic dome constructed with an aluminum framework is a nearly perfect hemisphere; its diameter measures 55.0 m on a winter day at a temperature of −15°C. How much more interior space does the dome have in the summer, when the temperature is 35°C?

17.17 •• A glass flask whose volume is 1000.00 cm^3 at 0.0°C is completely filled with mercury at this temperature. When flask and mercury are warmed to 55.0°C, 8.95 cm^3 of mercury overflow. If the coefficient of volume expansion of mercury is $18.0 \times 10^{-5}\ \text{K}^{-1}$, compute the coefficient of volume expansion of the glass.

17.18 •• A steel tank is completely filled with 1.90 m^3 of ethanol when both the tank and the ethanol are at 32.0°C. When the tank and its contents have cooled to 18.0°C, what additional volume of ethanol can be put into the tank?

17.19 •• A machinist bores a hole of diameter 1.35 cm in a steel plate that is at 25.0°C. What is the cross-sectional area of the hole (a) at 25.0°C and (b) when the temperature of the plate is increased to 175°C? Assume that the coefficient of linear expansion remains constant over this temperature range.

17.20 • Consider a flat metal plate with width w and length l, so its area is $A = lw$. The metal has coefficient of linear expansion α. Derive an expression, in terms of α, that gives the change ΔA in area for a change ΔT in temperature.

17.21 •• Steel train rails are laid in 12.0-m-long segments placed end to end. The rails are laid on a winter day when their temperature is −9.0°C. (a) How much space must be left between adjacent rails if they are just to touch on a summer day when their temperature is 33.0°C? (b) If the rails are originally laid in contact, what is the stress in them on a summer day when their temperature is 33.0°C?

17.22 •• A brass rod is 185 cm long and 1.60 cm in diameter. What force must be applied to each end of the rod to prevent it from contracting when it is cooled from 120.0°C to 10.0°C?

17.23 • The increase in length of an aluminum rod is twice the increase in length of an Invar rod with only a third of the temperature increase. Find the ratio of the lengths of the two rods.

Section 17.5 Quantity of Heat

17.24 • In an effort to stay awake for an all-night study session, a student makes a cup of coffee by first placing a 200 W electric immersion heater in 0.320 kg of water. (a) How much heat must be added to the water to raise its temperature from 20.0°C to 80.0°C? (b) How much time is required? Assume that all of the heater's power goes into heating the water.

17.25 •• An aluminum tea kettle with mass 1.10 kg and containing 1.80 kg of water is placed on a stove. If no heat is lost to the surroundings, how much heat must be added to raise the temperature from 20.0°C to 85.0°C?

17.26 • **BIO** **Heat Loss During Breathing.** In very cold weather a significant mechanism for heat loss by the human body is energy expended in warming the air taken into the lungs with each breath. (a) On a cold winter day when the temperature is $-20°C$, what amount of heat is needed to warm to body temperature ($37°C$) the 0.50 L of air exchanged with each breath? Assume that the specific heat of air is $1020\ J/kg \cdot K$ and that 1.0 L of air has mass 1.3×10^{-3} kg. (b) How much heat is lost per hour if the respiration rate is 20 breaths per minute?

17.27 • **BIO** While running, a 70 kg student generates thermal energy at a rate of 1200 W. For the runner to maintain a constant body temperature of 37°C, this energy must be removed by perspiration or other mechanisms. If these mechanisms failed and the energy could not flow out of the student's body, for what amount of time could a student run before irreversible body damage occurred? (*Note:* Protein structures in the body are irreversibly damaged if body temperature rises to 44°C or higher. The specific heat of a typical human body is $3480\ J/kg \cdot K$, slightly less than that of water. The difference is due to the presence of protein, fat, and minerals, which have lower specific heats.)

17.28 •• **On-Demand Water Heaters.** Conventional hot-water heaters consist of a tank of water maintained at a fixed temperature. The hot water is to be used when needed. The drawbacks are that energy is wasted because the tank loses heat when it is not in use and that you can run out of hot water if you use too much. Some utility companies are encouraging the use of *on-demand* water heaters (also known as *flash heaters*), which consist of heating units to heat the water as you use it. No water tank is involved, so no heat is wasted. A typical household shower flow rate is 2.5 gal/min (9.46 L/min) with the tap water being heated from 50°F (10°C) to 120°F (49°C) by the on-demand heater. What rate of heat input (either electrical or from gas) is required to operate such a unit, assuming that all the heat goes into the water?

17.29 • You are given a sample of metal and asked to determine its specific heat. You weigh the sample and find that its weight is 28.4 N. You carefully add 1.25×10^4 J of heat energy to the sample and find that its temperature rises 18.0 C°. What is the sample's specific heat?

17.30 • **CP** A 25,000 kg subway train initially traveling at 15.5 m/s slows to a stop in a station and then stays there long enough for its brakes to cool. The station's dimensions are 65.0 m long by 20.0 m wide by 12.0 m high. Assuming all the work done by the brakes in stopping the train is transferred as heat uniformly to all the air in the station, by how much does the air temperature in the station rise? Take the density of the air to be $1.20\ kg/m^3$ and its specific heat to be $1020\ J/kg \cdot K$.

17.31 • **CP** While painting the top of an antenna 225 m in height, a worker accidentally lets a 1.00 L water bottle fall from his lunchbox. The bottle lands in some bushes at ground level and does not break. If a quantity of heat equal to the magnitude of the change in mechanical energy of the water goes into the water, what is its increase in temperature?

17.32 • **CP** A nail driven into a board increases in temperature. If we assume that 60% of the kinetic energy delivered by a 1.80 kg hammer with a speed of 7.80 m/s is transformed into heat that flows into the nail and does not flow out, what is the temperature increase of an 8.00 g aluminum nail after it is struck ten times?

17.33 •• **CP** A 15.0 g bullet traveling horizontally at 865 m/s passes through a tank containing 13.5 kg of water and emerges with a speed of 534 m/s. What is the maximum temperature increase that the water could have as a result of this event?

Section 17.6 Calorimetry and Phase Changes

17.34 • You have 750 g of water at 10.0°C in a large insulated beaker. How much boiling water at 100.0°C must you add to this beaker so that the final temperature of the mixture will be 75°C?

17.35 •• A 500.0 g chunk of an unknown metal, which has been in boiling water for several minutes, is quickly dropped into an insulating Styrofoam beaker containing 1.00 kg of water at room temperature (20.0°C). After waiting and gently stirring for 5.00 minutes, you observe that the water's temperature has reached a constant value of 22.0°C. (a) Assuming that the Styrofoam absorbs a negligibly small amount of heat and that no heat was lost to the surroundings, what is the specific heat of the metal? (b) Which is more useful for storing thermal energy: this metal or an equal weight of water? Explain. (c) If the heat absorbed by the Styrofoam actually is not negligible, how would the specific heat you calculated in part (a) be in error? Would it be too large, too small, or still correct? Explain.

17.36 • **BIO** **Treatment for a Stroke.** One suggested treatment for a person who has suffered a stroke is immersion in an ice-water bath at 0°C to lower the body temperature, which prevents damage to the brain. In one set of tests, patients were cooled until their internal temperature reached 32.0°C. To treat a 70.0 kg patient, what is the minimum amount of ice (at 0°C) you need in the bath so that its temperature remains at 0°C? The specific heat of the human body is $3480\ J/kg \cdot C°$, and recall that normal body temperature is 37.0°C.

17.37 •• A blacksmith cools a 1.20 kg chunk of iron, initially at 650.0°C, by trickling 15.0°C water over it. All of the water boils away, and the iron ends up at 120.0°C. How much water did the blacksmith trickle over the iron?

17.38 •• A copper calorimeter can with mass 0.100 kg contains 0.160 kg of water and 0.0180 kg of ice in thermal equilibrium at atmospheric pressure. If 0.750 kg of lead at 255°C is dropped into the calorimeter can, what is the final temperature? Assume that no heat is lost to the surroundings.

17.39 •• A copper pot with a mass of 0.500 kg contains 0.170 kg of water, and both are at 20.0°C. A 0.250 kg block of iron at 85.0°C is dropped into the pot. Find the final temperature of the system, assuming no heat loss to the surroundings.

17.40 • In a container of negligible mass, 0.200 kg of ice at an initial temperature of $-40.0°C$ is mixed with a mass m of water that has an initial temperature of 80.0°C. No heat is lost to the surroundings. If the final temperature of the system is 28.0°C, what is the mass m of the water that was initially at 80.0°C?

17.41 • A 6.00 kg piece of solid copper metal at an initial temperature T is placed with 2.00 kg of ice that is initially at $-20.0°C$. The ice is in an insulated container of negligible mass and no heat is exchanged with the surroundings. After thermal equilibrium is reached, there is 1.20 kg of ice and 0.80 kg of liquid water. What was the initial temperature of the piece of copper?

17.42 •• An ice-cube tray of negligible mass contains 0.290 kg of water at 18.0°C. How much heat must be removed to cool the water to 0.00°C and freeze it? Express your answer in joules, calories, and Btu.

17.43 • How much heat is required to convert 18.0 g of ice at $-10.0°C$ to steam at 100.0°C? Express your answer in joules, calories, and Btu.

17.44 •• An open container holds 0.550 kg of ice at $-15.0°C$. The mass of the container can be ignored. Heat is supplied to the container at the constant rate of 800.0 J/min for 500.0 min. (a) After how many minutes does the ice *start* to melt? (b) After how many minutes, from the time when the heating is first started, does the temperature begin to rise above 0.0°C? (c) Plot a curve showing the temperature as a function of the elapsed time.

17.45 • **CP** What must the initial speed of a lead bullet be at 25.0°C so that the heat developed when it is brought to rest will be just sufficient to melt it? Assume that all the initial mechanical energy of the bullet is converted to heat and that no heat flows from the bullet to its surroundings. (Typical rifles have muzzle speeds that exceed the speed of sound in air, which is 347 m/s at 25.0°C.)

17.46 •• **BIO** **Steam Burns Versus Water Burns.** What is the amount of heat input to your skin when it receives the heat released (a) by 25.0 g of steam initially at 100.0°C, when it is cooled to skin temperature (34.0°C)? (b) By 25.0 g of water initially at 100.0°C, when it is cooled to 34.0°C? (c) What does this tell you about the relative severity of burns from steam versus burns from hot water?

17.47 • **BIO** **"The Ship of the Desert."** Camels require very little water because they are able to tolerate relatively large changes in their body temperature. While humans keep their body temperatures constant to within one or two Celsius degrees, a dehydrated camel permits its body temperature to drop to 34.0°C overnight and rise to 40.0°C during the day. To see how effective this mechanism is for saving water, calculate how many liters of water a 400 kg camel would have to drink if it attempted to keep its body temperature at a constant 34.0°C by evaporation of sweat during the day (12 hours) instead of letting it rise to 40.0°C. (*Note:* The specific heat of a camel or other mammal is about the same as that of a typical human, 3480 J/kg · K. The heat of vaporization of water at 34°C is 2.42×10^6 J/kg.)

17.48 • **BIO** Evaporation of sweat is an important mechanism for temperature control in some warm-blooded animals. (a) What mass of water must evaporate from the skin of a 70.0 kg man to cool his body 1.00 C°? The heat of vaporization of water at body temperature (37°C) is 2.42×10^6 J/kg. The specific heat of a typical human body is 3480 J/kg · K (see Exercise 17.27). (b) What volume of water must the man drink to replenish the evaporated water? Compare to the volume of a soft-drink can (355 cm^3).

17.49 •• **CP** An asteroid with a diameter of 10 km and a mass of 2.60×10^{15} kg impacts the earth at a speed of 32.0 km/s, landing in the Pacific Ocean. If 1.00% of the asteroid's kinetic energy goes to boiling the ocean water (assume an initial water temperature of 10.0°C), what mass of water will be boiled away by the collision? (For comparison, the mass of water contained in Lake Superior is about 2×10^{15} kg.)

17.50 • A laboratory technician drops a 0.0850 kg sample of unknown solid material, at 100.0°C, into a calorimeter. The calorimeter can, initially at 19.0°C, is made of 0.150 kg of copper and contains 0.200 kg of water. The final temperature of the calorimeter can and contents is 26.1°C. Compute the specific heat of the sample.

17.51 •• An insulated beaker with negligible mass contains 0.250 kg of water at 75.0°C. How many kilograms of ice at −20.0°C must be dropped into the water to make the final temperature of the system 40.0°C?

17.52 • A 4.00 kg silver ingot is taken from a furnace at 750.0°C and placed on a large block of ice at 0.0°C. Assuming that all the heat given up by the silver is used to melt the ice, how much ice is melted?

17.53 •• A plastic cup of negligible mass contains 0.280 kg of an unknown liquid at a temperature of 30.0°C. A 0.0270 kg mass of ice at a temperature of 0.0°C is added to the liquid, and when thermal equilibrium is reached the temperature of the combined substances is 14.0°C. Assuming no heat is exchanged with the surroundings, what is the specific heat capacity of the unknown liquid?

Section 17.7 Mechanisms of Heat Transfer

17.54 •• Two rods, one made of brass and the other made of copper, are joined end to end. The length of the brass section is 0.300 m and the length of the copper section is 0.800 m. Each segment has cross-sectional area 0.00500 m^2. The free end of the brass segment is in boiling water and the free end of the copper segment is in an ice–water mixture, in both cases under normal atmospheric pressure. The sides of the rods are insulated so there is no heat loss to the surroundings. (a) What is the temperature of the point where the brass and copper segments are joined? (b) What mass of ice is melted in 5.00 min by the heat conducted by the composite rod?

17.55 •• A copper bar is welded end to end to a bar of an unknown metal. The two bars have the same lengths and cross-sectional areas. The free end of the copper bar is maintained at a temperature T_H that can be varied. The free end of the unknown metal is kept at 0.0°C. To measure the thermal conductivity of the unknown metal, you measure the temperature T at the junction between the two bars for several values of T_H. You plot your data as T versus T_H, both in kelvins, and find that your data are well fit by a straight line that has slope 0.710. What do your measurements give for the value of the thermal conductivity of the unknown metal?

17.56 •• One end of an insulated metal rod is maintained at 100.0°C, and the other end is maintained at 0.00°C by an ice–water mixture. The rod is 60.0 cm long and has a cross-sectional area of 1.25 cm^2. The heat conducted by the rod melts 8.50 g of ice in 10.0 min. Find the thermal conductivity k of the metal.

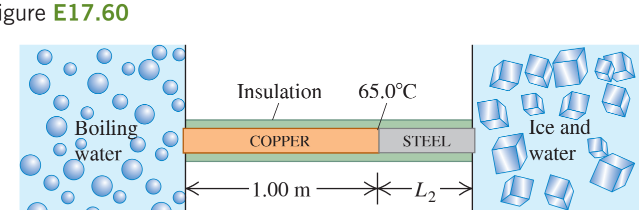
17.57 •• A carpenter builds an exterior house wall with a layer of wood 3.0 cm thick on the outside and a layer of Styrofoam insulation 2.2 cm thick on the inside wall surface. The wood has $k = 0.080$ W/m · K, and the Styrofoam has $k = 0.027$ W/m · K. The interior surface temperature is 19.0°C, and the exterior surface temperature is −10.0°C. (a) What is the temperature at the plane where the wood meets the Styrofoam? (b) What is the rate of heat flow per square meter through this wall?

17.58 • An electric kitchen range has a total wall area of 1.40 m^2 and is insulated with a layer of fiberglass 4.00 cm thick. The inside surface of the fiberglass has a temperature of 175°C, and its outside surface is at 35.0°C. The fiberglass has a thermal conductivity of 0.040 W/m · K. (a) What is the heat current through the insulation, assuming it may be treated as a flat slab with an area of 1.40 m^2? (b) What electric-power input to the heating element is required to maintain this temperature?

17.59 • Air has a very low thermal conductivity. This explains why we feel comfortable wearing short sleeves in a 20°C environment even though our body temperature is 37°C. Material objects feel cooler to our immediate touch than the air, owing to relatively high thermal conductivities. (a) Touch a few surfaces in a room-temperature environment and rank them in order of which feel the coolest to which feel the warmest. Objects that feel cooler have larger thermal conductivities. Consider a wood surface, a metallic surface, and a glass surface, and rank these in order from coolest to warmest. (b) How does your ranking compare to the thermal conductivities listed in Table 17.5?

17.60 • A long rod, insulated to prevent heat loss along its sides, is in perfect thermal contact with boiling water (at atmospheric pressure) at one end and with an ice–water mixture at the other (**Fig. E17.60**). The rod consists of a 1.00 m section of copper (one end in boiling water) joined end to end to a length L_2 of steel (one end in the ice–water mixture). Both sections of the rod have cross-sectional areas of 4.00 cm^2. The temperature of the copper–steel junction is 65.0°C after a steady state has been set up. (a) How much heat per second flows from the boiling water to the ice–water mixture? (b) What is the length L_2 of the steel section?

Figure **E17.60**



17.61 • A pot with a steel bottom 8.50 mm thick rests on a hot stove. The area of the bottom of the pot is $0.150\ \text{m}^2$. The water inside the pot is at 100.0°C, and 0.390 kg are evaporated every 3.00 min. Find the temperature of the lower surface of the pot, which is in contact with the stove.

17.62 •• You are asked to design a cylindrical steel rod 50.0 cm long, with a circular cross section, that will conduct 190.0 J/s from a furnace at 400.0°C to a container of boiling water under 1 atmosphere. What must the rod's diameter be?

17.63 •• A picture window has dimensions of $1.40\ \text{m} \times 2.50\ \text{m}$ and is made of glass 5.20 mm thick. On a winter day, the temperature of the outside surface of the glass is -20.0°C, while the temperature of the inside surface is a comfortable 19.5°C. (a) At what rate is heat being lost through the window by conduction? (b) At what rate would heat be lost through the window if you covered it with a 0.750-mm-thick layer of paper (thermal conductivity $0.0500\ \text{W/m} \cdot \text{K}$)?

17.64 • What is the rate of energy radiation per unit area of a blackbody at (a) 273 K and (b) 2730 K?

17.65 • **Size of a Light-Bulb Filament.** The operating temperature of a tungsten filament in an incandescent light bulb is 2450 K, and its emissivity is 0.350. Find the surface area of the filament of a 150 W bulb if all the electrical energy consumed by the bulb is radiated by the filament as electromagnetic waves. (Only a fraction of the radiation appears as visible light.)

17.66 •• The emissivity of tungsten is 0.350. A tungsten sphere with radius 1.50 cm is suspended within a large evacuated enclosure whose walls are at 290.0 K. What power input is required to maintain the sphere at 3000.0 K if heat conduction along the supports is ignored?

17.67 • **The Sizes of Stars.** The hot glowing surfaces of stars emit energy in the form of electromagnetic radiation. It is a good approximation to assume $e = 1$ for these surfaces. Find the radii of the following stars (assumed to be spherical): (a) Rigel, the bright blue star in the constellation Orion, which radiates energy at a rate of 2.7×10^{32} W and has surface temperature 11,000 K; (b) Procyon B (visible only using a telescope), which radiates energy at a rate of 2.1×10^{23} W and has surface temperature 10,000 K. (c) Compare your answers to the radius of the earth, the radius of the sun, and the distance between the earth and the sun. (Rigel is an example of a *supergiant* star, and Procyon B is an example of a *white dwarf* star.)

PROBLEMS

17.68 •• Figure 17.12 shows that the graph of the volume of 1 gram of liquid water can be closely approximated by a parabola in the temperature range between 0°C and 10°C. (a) Show that the equation of this parabola has the form $V = A + B(T_C - 4.0°\text{C})^2$ and find the values of the constants A and B. (b) Define the temperature-dependent quantity $\beta(T_C)$ in terms of the equation $dV = \beta(T_C)\, V\, dT$. Use the result of part (a) to find the value of $\beta(T_C)$ for $T_C = 1.0$°C, 4.0°C, 7.0°C, and 10.0°C. Your results show that β is not constant in this temperature range but is approximately constant above 7.0°C.

17.69 •• **CP** A Foucault pendulum consists of a brass sphere with a diameter of 35.0 cm suspended from a steel cable 10.5 m long (both measurements made at 20.0°C). Due to a design oversight, the swinging sphere clears the floor by a distance of only 2.00 mm when the temperature is 20.0°C. At what temperature will the sphere begin to brush the floor?

17.70 •• A steel wire has density $7800\ \text{kg/m}^3$ and mass 2.50 g. It is stretched between two rigid supports separated by 0.400 m. (a) When the temperature of the wire is 20.0°C, the frequency of the fundamental standing wave for the wire is 440 Hz. What is the tension in the wire? (b) What is the temperature of the wire if its fundamental standing wave has frequency 460 Hz? For steel the coefficient of linear expansion is $1.2 \times 10^{-5}\ \text{K}^{-1}$ and Young's modulus is 20×10^{10} Pa.

17.71 •• An unknown liquid has density ρ and coefficient of volume expansion β. A quantity of heat Q is added to a volume V of the liquid, and the volume of the liquid increases by an amount ΔV. There is no phase change. In terms of these quantities, what is the specific heat capacity c of the liquid?

17.72 •• **CP** A small fused quartz sphere swings back and forth as a simple pendulum on the lower end of a long copper wire that is attached to the ceiling at its upper end. The amplitude of swing is small. When the wire has a temperature of 20.0°C, its length is 3.00 m. What is the percentage change in the period of the motion if the temperature of the wire is increased to 220°C? (*Hint:* Use the power series expansion for $(1 + x)^n$ in Appendix B.)

17.73 ••• You propose a new temperature scale with temperatures given in °M. You define 0.0°M to be the normal melting point of mercury and 100.0°M to be the normal boiling point of mercury. (a) What is the normal boiling point of water in °M? (b) A temperature change of 10.0 M° corresponds to how many C°?

17.74 • **CP CALC** A 250 kg weight is hanging from the ceiling by a thin copper wire. In its fundamental mode, this wire vibrates at the frequency of concert A (440 Hz). You then increase the temperature of the wire by 40 C°. (a) By how much will the fundamental frequency change? Will it increase or decrease? (b) By what percentage will the speed of a wave on the wire change? (c) By what percentage will the wavelength of the fundamental standing wave change? Will it increase or decrease?

17.75 ••• You are making pesto for your pasta and have a cylindrical measuring cup 10.0 cm high made of ordinary glass $[\beta = 2.7 \times 10^{-5}\ (\text{C}°)^{-1}]$ that is filled with olive oil $[\beta = 6.8 \times 10^{-4}\ (\text{C}°)^{-1}]$ to a height of 3.00 mm below the top of the cup. Initially, the cup and oil are at room temperature (22.0°C). You get a phone call and forget about the olive oil, which you inadvertently leave on the hot stove. The cup and oil heat up slowly and have a common temperature. At what temperature will the olive oil start to spill out of the cup?

17.76 •• A surveyor's 30.0 m steel tape is correct at 20.0°C. The distance between two points, as measured by this tape on a day when its temperature is 5.00°C, is 25.970 m. What is the true distance between the points?

17.77 •• A metal rod that is 30.0 cm long expands by 0.0650 cm when its temperature is raised from 0.0°C to 100.0°C. A rod of a different metal and of the same length expands by 0.0350 cm for the same rise in temperature. A third rod, also 30.0 cm long, is made up of pieces of each of the above metals placed end to end and expands 0.0580 cm between 0.0°C and 100.0°C. Find the length of each portion of the composite rod.

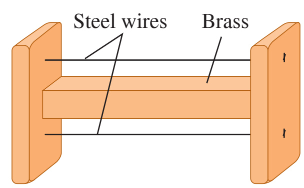
17.78 •• A copper sphere with density $8900\ \text{kg/m}^3$, radius 5.00 cm, and emissivity $e = 1.00$ sits on an insulated stand. The initial temperature of the sphere is 300 K. The surroundings are very cold, so the rate of absorption of heat by the sphere can be neglected. (a) How long does it take the sphere to cool by 1.00 K due to its radiation of heat energy? Neglect the change in heat current as the temperature decreases. (b) To assess the accuracy of the approximation used in part (a), what is the fractional change in the heat current H when the temperature changes from 300 K to 299 K?

17.79 ••• (a) Equation (17.12) gives the stress required to keep the length L of a rod constant as its temperature T changes. Show that if L is permitted to change by an amount ΔL when T changes by ΔT, the stress is

$$\frac{F}{A} = Y\left(\frac{\Delta L}{L_0} - \alpha\,\Delta T\right)$$

where F is the tension on the rod, L_0 is the original length of the rod, A its cross-sectional area, α its coefficient of linear expansion, and Y its Young's modulus. (b) A heavy brass bar has projections at its ends (**Fig. P17.79**). Two fine steel wires, fastened between the projections, are just taut (zero tension) when the whole system is at 20°C. What is the tensile stress in the steel wires when the temperature of the system is raised to 140°C? Make any simplifying assumptions you think are justified, but state them.

Figure **P17.79**



17.80 •• **CP** A metal wire, with density ρ and Young's modulus Y, is stretched between rigid supports. At temperature T, the speed of a transverse wave is found to be v_1. When the temperature is increased to $T + \Delta T$, the speed decreases to $v_2 < v_1$. Determine the coefficient of linear expansion of the wire.

17.81 •• A steel ring with a 2.5000 in. inside diameter at 20.0°C is to be warmed and slipped over a brass shaft with a 2.5020 in. outside diameter at 20.0°C. (a) To what temperature should the ring be warmed? (b) If the ring and the shaft together are cooled by some means such as liquid air, at what temperature will the ring just slip off the shaft?

17.82 • **BIO** **Doughnuts: Breakfast of Champions!** A typical doughnut contains 2.0 g of protein, 17.0 g of carbohydrates, and 7.0 g of fat. Average food energy values are 4.0 kcal/g for protein and carbohydrates and 9.0 kcal/g for fat. (a) During heavy exercise, an average person uses energy at a rate of 510 kcal/h. How long would you have to exercise to "work off" one doughnut? (b) If the energy in the doughnut could somehow be converted into the kinetic energy of your body as a whole, how fast could you move after eating the doughnut? Take your mass to be 60 kg, and express your answer in m/s and in km/h.

17.83 •• There is 0.050 kg of an unknown liquid in a plastic container of negligible mass. The liquid has a temperature of 90.0°C. To measure the specific heat capacity of the unknown liquid, you add a mass m_w of water that has a temperature of 0.0°C to the liquid and measure the final temperature T after the system has reached thermal equilibrium. You repeat this measurement for several values of m_w, with the initial temperature of the unknown liquid always equal to 90.0°C. The plastic container is insulated, so no heat is exchanged with the surroundings. You plot your data as m_w versus T^{-1}, the inverse of the final temperature T. Your data points lie close to a straight line that has slope 2.15 kg · C°. What does this result give for the value of the specific heat capacity of the unknown liquid?

17.84 •• You cool a 100.0 g slug of red-hot iron (temperature 745°C) by dropping it into an insulated cup of negligible mass containing 85.0 g of water at 20.0°C. Assuming no heat exchange with the surroundings, (a) what is the final temperature of the water and (b) what is the final mass of the iron and the remaining water?

17.85 •• **CALC** **Debye's T^3 Law.** At very low temperatures the molar heat capacity of rock salt varies with temperature according to Debye's T^3 law:

$$C = k\frac{T^3}{\theta^3}$$

where $k = 1940$ J/mol · K and $\theta = 281$ K. (a) How much heat is required to raise the temperature of 1.50 mol of rock salt from 10.0 K to 40.0 K? (*Hint:* Use Eq. (17.18) in the form $dQ = nC\,dT$ and integrate.) (b) What is the average molar heat capacity in this range? (c) What is the true molar heat capacity at 40.0 K?

17.86 •• The heat one feels when sitting near the fire in a fireplace or at a campfire is due almost entirely to thermal radiation. (a) Estimate the diameter and length of an average campfire log. (b) Compute the surface area of such a log. (c) Use the Stefan–Boltzmann law to determine the power emitted by thermal radiation by such a log when it burns at a typical temperature of 700°C in a surrounding air temperature of 20.0°C. The emissivity of a burning log is close to unity.

17.87 • **Hot Air in a Physics Lecture.** (a) A typical student listening attentively to a physics lecture has a heat output of 100 W. How much heat energy does a class of 140 physics students release into a lecture hall over the course of a 50 min lecture? (b) Assume that all the heat energy in part (a) is transferred to the 3200 m^3 of air in the room. The air has specific heat 1020 J/kg · K and density 1.20 kg/m^3. If none of the heat escapes and the air conditioning system is off, how much will the temperature of the air in the room rise during the 50 min lecture? (c) If the class is taking an exam, the heat output per student rises to 280 W. What is the temperature rise during 50 min in this case?

17.88 ••• **CALC** The molar heat capacity of a certain substance varies with temperature according to the empirical equation

$$C = 29.5 \text{ J/mol}\cdot\text{K} + (8.20 \times 10^{-3} \text{ J/mol}\cdot\text{K}^2)T$$

How much heat is necessary to change the temperature of 3.00 mol of this substance from 27°C to 227°C? (*Hint:* Use Eq. (17.18) in the form $dQ = nC\,dT$ and integrate.)

17.89 •• **BIO** **Bicycling on a Warm Day.** If the air temperature is the same as the temperature of your skin (about 30°C), your body cannot get rid of heat by transferring it to the air. In that case, it gets rid of the heat by evaporating water (sweat). During bicycling, a typical 70 kg person's body produces energy at a rate of about 500 W due to metabolism, 80% of which is converted to heat. (a) How many kilograms of water must the person's body evaporate in an hour to get rid of this heat? The heat of vaporization of water at body temperature is 2.42×10^6 J/kg. (b) The evaporated water must, of course, be replenished, or the person will dehydrate. How many 750 mL bottles of water must the bicyclist drink per hour to replenish the lost water? (Recall that the mass of a liter of water is 1.0 kg.)

17.90 •• **BIO** **Overheating.** (a) By how much would the body temperature of the bicyclist in Problem 17.89 increase in an hour if he were unable to get rid of the excess heat? (b) Is this temperature increase large enough to be serious? To find out, how high a fever would it be equivalent to, in °F? (Recall that the normal internal body temperature is 98.6°F and the specific heat of the body is 3480 J/kg · C°.)

17.91 • **BIO** **A Thermodynamic Process in an Insect.** The African bombardier beetle (*Stenaptinus insignis*) can emit a jet of defensive spray from the movable tip of its abdomen (**Fig. P17.91**). The beetle's body has reservoirs containing two chemicals; when the beetle is disturbed, these chemicals combine in a reaction chamber, producing a compound that is warmed from 20°C to 100°C by the heat of reaction. The high pressure produced allows the compound to be sprayed out at speeds up to 19 m/s (68 km/h), scaring away predators of all kinds. (The beetle shown in Fig. P17.91 is 2 cm long.) Calculate the heat of reaction of the two chemicals (in J/kg). Assume that the specific heat of the chemicals and of the spray is the same as that of water, 4.19×10^3 J/kg · K, and that the initial temperature of the chemicals is 20°C.

Figure **P17.91**

17.92 ••• **CP** A industrious explorer of the polar regions has devised a contraption for melting ice. It consists of a sealed 10 L cylindrical tank with a porous grate separating the top half from the bottom half. The bottom half includes a paddle wheel attached to an axle that passes outside the cylinder, where it is attached by a gearbox and pulley system to a stationary bicycle. Pedaling the bicycle rotates the paddle wheel inside the cylinder. The tank includes 6.00 L of water and 3.00 kg of ice at 0.0°C. The water fills the bottom chamber, where it may be agitated by the paddle wheel, and partially fills the upper chamber, which also includes the ice. The bicycle is pedaled with an average torque of 25.0 N·m at a rate of 30.0 revolutions per minute. The system is 70% efficient. (a) For what length of time must the explorer pedal the bicycle to melt all the ice? (b) How much longer must he pedal to raise the temperature of the water to 10.5°C?

17.93 •• You have 1.50 kg of water at 28.0°C in an insulated container of negligible mass. You add 0.600 kg of ice that is initially at −22.0°C. Assume that no heat exchanges with the surroundings. (a) After thermal equilibrium has been reached, has all of the ice melted? (b) If all of the ice has melted, what is the final temperature of the water in the container? If some ice remains, what is the final temperature of the water in the container, and how much ice remains?

17.94 •• A thirsty nurse cools a 2.00 L bottle of a soft drink (mostly water) by pouring it into a large aluminum mug of mass 0.257 kg and adding 0.120 kg of ice initially at −15.0°C. If the soft drink and mug are initially at 20.0°C, what is the final temperature of the system, assuming that no heat is lost?

17.95 ••• A copper calorimeter can with mass 0.446 kg contains 0.0950 kg of ice. The system is initially at 0.0°C. (a) If 0.0350 kg of steam at 100.0°C and 1.00 atm pressure is added to the can, what is the final temperature of the calorimeter can and its contents? (b) At the final temperature, how many kilograms are there of ice, how many of liquid water, and how many of steam?

17.96 • A Styrofoam bucket of negligible mass contains 1.75 kg of water and 0.450 kg of ice. More ice, from a refrigerator at −15.0°C, is added to the mixture in the bucket, and when thermal equilibrium has been reached, the total mass of ice in the bucket is 0.884 kg. Assuming no heat exchange with the surroundings, what mass of ice was added?

17.97 ••• In a container of negligible mass, 0.0400 kg of steam at 100°C and atmospheric pressure is added to 0.200 kg of water at 50.0°C. (a) If no heat is lost to the surroundings, what is the final temperature of the system? (b) At the final temperature, how many kilograms are there of steam and how many of liquid water?

17.98 •• **BIO** **Mammal Insulation.** Animals in cold climates often depend on *two* layers of insulation: a layer of body fat (of thermal conductivity 0.20 W/m·K) surrounded by a layer of air trapped inside fur or down. We can model a black bear (*Ursus americanus*) as a sphere 1.5 m in diameter having a layer of fat 4.0 cm thick. (Actually, the thickness varies with the season, but we are interested in hibernation, when the fat layer is thickest.) In studies of bear hibernation, it was found that the outer surface layer of the fur is at 2.7°C during hibernation, while the inner surface of the fat layer is at 31.0°C. (a) What is the temperature at the fat–inner fur boundary so that the bear loses heat at a rate of 50.0 W? (b) How thick should the air layer (contained within the fur) be?

17.99 •• **Effect of a Window in a Door.** A carpenter builds a solid wood door with dimensions 2.00 m × 0.95 m × 5.0 cm. Its thermal conductivity is $k = 0.120$ W/m·K. The air films on the inner and outer surfaces of the door have the same combined thermal resistance as an additional 1.8 cm thickness of solid wood. The inside air temperature is 20.0°C, and the outside air temperature is −8.0°C. (a) What is the rate of heat flow through the door? (b) By what factor is the heat flow increased if a window 0.500 m on a side is inserted in the door? The glass is 0.450 cm thick, and the glass has a thermal conductivity of 0.80 W/m·K. The air films on the two sides of the glass have a total thermal resistance that is the same as an additional 12.0 cm of glass.

17.100 ••• **CP** At 0°C, a cylindrical metal bar with radius r and mass M is slid snugly into a circular hole in a large, horizontal, rigid slab of thickness d. For this metal, Young's modulus is Y and the coefficient of linear expansion is α. A light but strong hook is attached to the underside of the metal bar; this apparatus is used as part of a hoist in a shipping yard. The coefficient of static friction between the bar and the slab is μ_s. At a temperature T above 0°C, the hook is attached to a large container and the slab is raised. What is the largest mass the container can have without the metal bar slipping out of the slab as the container is slowly lifted? The slab undergoes negligible thermal expansion.

17.101 •• Compute the ratio of the rate of heat loss through a single-pane window with area 0.15 m^2 to that for a double-pane window with the same area. The glass of a single pane is 4.2 mm thick, and the air space between the two panes of the double-pane window is 7.0 mm thick. The glass has thermal conductivity 0.80 W/m·K. The air films on the room and outdoor surfaces of either window have a combined thermal resistance of 0.15 m^2·K/W.

17.102 • Rods of copper, brass, and steel—each with cross-sectional area of 2.00 cm^2—are welded together to form a Y-shaped figure. The free end of the copper rod is maintained at 100.0°C, and the free ends of the brass and steel rods at 0.0°C. Assume that there is no heat loss from the surfaces of the rods. The lengths of the rods are: copper, 13.0 cm; brass, 18.0 cm; steel, 24.0 cm. What is (a) the temperature of the junction point; (b) the heat current in each of the three rods?

17.103 •• **BIO** **Jogging in the Heat of the Day.** You have probably seen people jogging in extremely hot weather. There are good reasons not to do this! When jogging strenuously, an average runner of mass 68 kg and surface area 1.85 m^2 produces energy at a rate of up to 1300 W, 80% of which is converted to heat. The jogger radiates heat but actually absorbs more from the hot air than he radiates away. At such high levels of activity, the skin's temperature can be elevated to around 33°C instead of the usual 30°C. (Ignore conduction, which would bring even more heat into his body.) The only way for the body to get rid of this extra heat is by evaporating water (sweating). (a) How much heat per second is produced just by the act of jogging? (b) How much *net* heat per second does the runner gain just from radiation if the air temperature is 40.0°C (104°F)? (Remember: He radiates out, but the environment radiates back in.) (c) What is the *total* amount of excess heat this runner's body must get rid of per second? (d) How much water must his body evaporate every minute due to his activity? The heat of vaporization of water at body temperature is 2.42×10^6 J/kg. (e) How many 750 mL bottles of water must he drink after (or preferably before!) jogging for a half hour? Recall that a liter of water has a mass of 1.0 kg.

17.104 ••• **BIO** **Basal Metabolic Rate.** The *basal metabolic rate* is the rate at which energy is produced in the body when a person is at rest. A 75 kg (165 lb) person of height 1.83 m (6 ft) has a body surface area of approximately 2.0 m^2. (a) What is the net amount of heat this person could radiate per second into a room at 18°C (about 65°F) if his skin's surface temperature is 30°C? (At such temperatures, nearly all the heat is infrared radiation, for which the body's emissivity is 1.0, regardless of the amount of pigment.) (b) Normally, 80% of the energy produced by metabolism goes into heat, while the rest goes into things like pumping blood and repairing cells. Also normally, a person at rest can get rid of this excess heat just through radiation. Use your answer to part (a) to find this person's basal metabolic rate.

17.105 ••• CALC Time Needed for a Lake to Freeze Over. (a) When the air temperature is below 0°C, the water at the surface of a lake freezes to form an ice sheet. Why doesn't freezing occur throughout the entire volume of the lake? (b) Show that the thickness of the ice sheet formed on the surface of a lake is proportional to the square root of the time if the heat of fusion of the water freezing on the underside of the ice sheet is conducted through the sheet. (c) Assuming that the upper surface of the ice sheet is at -10°C and the bottom surface is at 0°C, calculate the time it will take to form an ice sheet 25 cm thick. (d) If the lake in part (c) is uniformly 40 m deep, how long would it take to freeze all the water in the lake? Is this likely to occur?

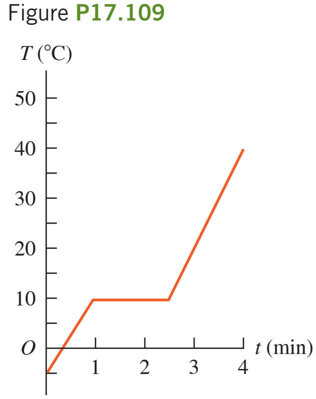
17.106 • The rate at which radiant energy from the sun reaches the earth's upper atmosphere is about $1.50\ \text{kW/m}^2$. The distance from the earth to the sun is 1.50×10^{11} m, and the radius of the sun is 6.96×10^8 m. (a) What is the rate of radiation of energy per unit area from the sun's surface? (b) If the sun radiates as an ideal blackbody, what is the temperature of its surface?

17.107 ••• A Thermos for Liquid Helium. A physicist uses a cylindrical metal can 0.250 m high and 0.090 m in diameter to store liquid helium at 4.22 K; at that temperature the heat of vaporization of helium is 2.09×10^4 J/kg. Completely surrounding the metal can are walls maintained at the temperature of liquid nitrogen, 77.3 K, with vacuum between the can and walls. How much liquid helium boils away per hour? The emissivity of the metal can is 0.200. The only heat transfer between the metal can and the surrounding walls is by radiation.

17.108 •• A metal sphere with radius 3.20 cm is suspended in a large metal box with interior walls that are maintained at 30.0°C. A small electric heater is embedded in the sphere. Heat energy must be supplied to the sphere at the rate of 0.660 J/s to maintain the sphere at a constant temperature of 41.0°C. (a) What is the emissivity of the metal sphere? (b) What power input to the sphere is required to maintain it at 82.0°C? What is the ratio of the power required for 82.0°C to the power required for 41.0°C? How does this ratio compare with 2^4? Explain.

17.109 •• DATA As a physicist, you put heat into a 500.0 g solid sample at the rate of 10.0 kJ/min while recording its temperature as a function of time. You plot your data as shown in **Fig. P17.109**. (a) What is the latent heat of fusion for this solid? (b) What are the specific heats of the liquid and solid states of this material?

Figure P17.109



17.110 ••• DATA At a chemical plant where you are an engineer, a tank contains an unknown liquid. You must determine the liquid's specific heat capacity. You put 0.500 kg of the liquid into an insulated metal cup of mass 0.200 kg. Initially the liquid and cup are at 20.0°C. You add 0.500 kg of water that has a temperature of 80.0°C. After thermal equilibrium has been reached, the final temperature of the two liquids and the cup is 58.1°C. You then empty the cup and repeat the experiment with the same initial temperatures, but this time with 1.00 kg of the unknown liquid. The final temperature is 49.3°C. Assume that the specific heat capacities are constant over the temperature range of the experiment and that no heat is lost to the surroundings. Calculate the specific heat capacity of the liquid and of the metal from which the cup is made.

17.111 •• DATA As a mechanical engineer, you are given two uniform metal bars A and B, made from different metals, to determine their thermal conductivities. You measure both bars to have length 40.0 cm and uniform cross-sectional area $2.50\ \text{cm}^2$. You place one end of bar A in thermal contact with a very large vat of boiling water at 100.0°C and the other end in thermal contact with an ice–water mixture at 0.0°C. To prevent heat loss along the bar's sides, you wrap insulation around the bar. You weigh the amount of ice initially as 300 g. After 45.0 min, you weigh the ice again; 191 g of ice remains. The ice–water mixture is in an insulated container, so the only heat entering or leaving it is the heat conducted by the metal bar.

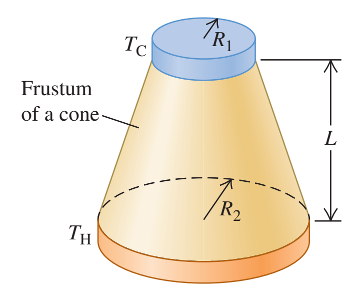
You are confident that your data will allow you to calculate the thermal conductivity k_A of bar A. But this measurement was tedious—you don't want to repeat it for bar B. Instead, you glue the bars together end to end, with adhesive that has very large thermal conductivity, to make a composite bar 80.0 m long. You place the free end of A in thermal contact with the boiling water and the free end of B in thermal contact with the ice–water mixture. The composite bar is thermally insulated. Hours later, you notice that ice remains in the ice–water mixture. Measuring the temperature at the junction of the two bars, you find that it is 62.4°C. After 10 min you repeat that measurement and get the same temperature, with ice remaining in the ice–water mixture. From your data, calculate the thermal conductivities of bar A and of bar B.

CHALLENGE PROBLEMS

17.112 • At a remote arctic research base, liquid water is obtained by melting ice in a propane-fueled conversion tank. Propane has a heat of combustion of 25.6 MJ/L, and 30% of the released energy supplies heat to the tank. Liquid water at 0°C is drawn off the tank at a rate of 500 mL/min, while a corresponding amount of ice at 0°C is continually inserted into the tank from a hopper. How long will an 18 L tank of propane fuel this operation?

17.113 ••• CALC A frustum of a cone (**Fig. P17.113**) has smaller radius R_1, larger radius R_2, and length L and is made from a material with thermal conductivity k. Derive an expression for the conductive heat current through the frustum when the side with radius R_1 is kept at temperature T_H and the side with radius R_2 is kept at temperature T_C. [*Hint:* Parameterize the axis of the frustum using coordinate x. Use Eq. (17.21) for the heat current H through a differential slice of the frustum with length dx, area $A = \pi r^2$ (where r is a function of x), and temperature difference dT. Separate variables and integrate on dT and dx.]

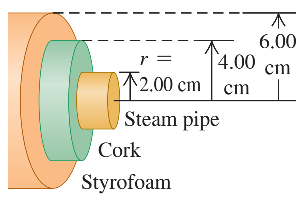
Figure P17.113



17.114 ••• BIO A Walk in the Sun. Consider a poor lost soul walking at 5 km/h on a hot day in the desert, wearing only a bathing suit. This person's skin temperature tends to rise due to four mechanisms: (i) energy is generated by metabolic reactions in the body at a rate of 280 W, and almost all of this energy is converted to heat that flows to the skin; (ii) heat is delivered to the skin by convection from the outside air at a rate equal to $k'A_{\text{skin}}(T_{\text{air}} - T_{\text{skin}})$, where k' is $54\ \text{J/h} \cdot \text{C}° \cdot \text{m}^2$, the exposed skin area A_{skin} is $1.5\ \text{m}^2$, the air temperature T_{air} is 47°C, and the skin temperature T_{skin} is 36°C; (iii) the skin absorbs radiant energy from the sun at a rate of $1400\ \text{W/m}^2$; (iv) the skin absorbs radiant energy from the environment, which has temperature 47°C. (a) Calculate the net rate (in watts) at which the person's skin is heated by all four of these mechanisms. Assume that the emissivity of the skin is $e = 1$ and that the skin temperature is initially 36°C. Which mechanism is the most important? (b) At what rate (in L/h) must perspiration evaporate from this person's skin to maintain a constant skin temperature? (The heat of vaporization of water at 36°C is 2.42×10^6 J/kg.) (c) Suppose the person is protected by light-colored clothing ($e \approx 0$) and only $0.45\ \text{m}^2$ of skin is exposed. What rate of perspiration is required now? Discuss the usefulness of the traditional clothing worn by desert peoples.

17.115 ••• A hollow cylinder has length L, inner radius a, and outer radius b, and the temperatures at the inner and outer surfaces are T_2 and T_1. (The cylinder could represent an insulated hot-water pipe.) The thermal conductivity of the material of which the cylinder is made is k. Derive an equation for (a) the total heat current through the walls of the cylinder; (b) the temperature variation inside the cylinder walls. (c) Show that the equation for the total heat current reduces to Eq. (17.21) for linear heat flow when the cylinder wall is very thin. (d) A steam pipe with a radius of 2.00 cm, carrying steam at 140°C, is surrounded by a cylindrical jacket with inner and outer radii 2.00 cm and 4.00 cm and made of a type of cork with thermal conductivity 4.00×10^{-2} W/m·K. This in turn is surrounded by a cylindrical jacket made of a brand of Styrofoam with thermal conductivity 2.70×10^{-2} W/m·K and having inner and outer radii 4.00 cm and 6.00 cm (**Fig. P17.115**). The outer surface of the Styrofoam has a temperature of 15°C. What is the temperature at a radius of 4.00 cm, where the two insulating layers meet? (e) What is the total rate of transfer of heat out of a 2.00 m length of pipe?

Figure **P17.115**



MCAT-STYLE PASSAGE PROBLEMS

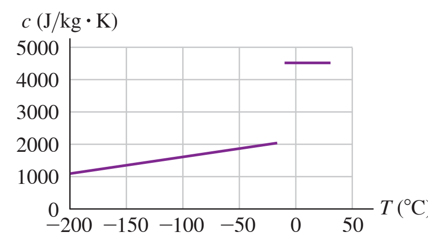
BIO Preserving Cells at Cold Temperatures. In cryopreservation, biological materials are cooled to a very low temperature to slow down chemical reactions that might damage the cells or tissues. It is important to prevent the materials from forming ice crystals during freezing. One method for preventing ice formation is to place the material in a protective solution called a *cryoprotectant*. Stated values of the thermal properties of one cryoprotectant are listed here:

Melting point	−20°C
Latent heat of fusion	2.80×10^5 J/kg
Specific heat (liquid)	4.5×10^3 J/kg·K
Specific heat (solid)	2.0×10^3 J/kg·K
Thermal conductivity (liquid)	1.2 W/m·K
Thermal conductivity (solid)	2.5 W/m·K

17.116 You place 35 g of this cryoprotectant at 22°C in contact with a cold plate that is maintained at the boiling temperature of liquid nitrogen (77 K). The cryoprotectant is thermally insulated from everything but the cold plate. Use the values in the table to determine how much heat will be transferred from the cryoprotectant as it reaches thermal equilibrium with the cold plate. (a) 1.5×10^4 J; (b) 2.9×10^4 J; (c) 3.4×10^4 J; (d) 4.4×10^4 J.

17.117 Careful measurements show that the specific heat of the solid phase depends on temperature (**Fig. P17.117**). How will the actual time needed for this cryoprotectant to come to equilibrium with the cold plate compare with the time predicted by using the values in the table? Assume that all values other than the specific heat (solid) are correct. The actual time (a) will be shorter; (b) will be longer; (c) will be the same; (d) depends on the density of the cryoprotectant.

Figure **P17.117**



17.118 In another experiment, you place a layer of this cryoprotectant between one 10 cm × 10 cm cold plate maintained at −40°C and a second cold plate of the same size maintained at liquid nitrogen's boiling temperature (77 K). Then you measure the rate of heat transfer. Another lab wants to repeat the experiment but uses cold plates that are 20 cm × 20 cm, with one at −40°C and the other at 77 K. How thick does the layer of cryoprotectant have to be so that the rate of heat transfer by conduction is the same as that when you use the smaller plates? (a) One-quarter the thickness; (b) half the thickness; (c) twice the thickness; (d) four times the thickness.

17.119 To measure the specific heat in the liquid phase of a newly developed cryoprotectant, you place a sample of the new cryoprotectant in contact with a cold plate until the solution's temperature drops from room temperature to its freezing point. Then you measure the heat transferred to the cold plate. If the system isn't sufficiently isolated from its room-temperature surroundings, what will be the effect on the measurement of the specific heat? (a) The measured specific heat will be greater than the actual specific heat; (b) the measured specific heat will be less than the actual specific heat; (c) there will be no effect because the thermal conductivity of the cryoprotectant is so low; (d) there will be no effect on the specific heat, but the temperature of the freezing point will change.

ANSWERS

Chapter Opening Question ?

(iii) The molten iron contains a large amount of energy. An object *has* a temperature but does not *contain* temperature. By "heat" we mean energy that is in transit from one object to another as a result of temperature difference between the objects. Objects do not *contain* heat.

Key Example VARIATION Problems

VP17.4.1 (a) $2.00 \times 10^{-5}\ \mathrm{K}^{-1}$ (b) length decreases by 0.270 mm

VP17.4.2 12 cm^3

VP17.4.3 1.7×10^3 N; tensile

VP17.4.4 $\dfrac{1}{\alpha_B - \alpha_A}\left(\alpha_B L - \dfrac{\Delta L}{\Delta T}\right)$

VP17.9.1 0.20 kg

VP17.9.2 14.0°C

VP17.9.3 0.181 kg

VP17.9.4 (a) 3.31°C (b) all of it

VP17.15.1 (a) 0.754 W/m·K (b) 733 W

VP17.15.2 (a) 254°C (b) −31°C

VP17.15.3 1.2×10^9 m; 1.7 times the sun's radius

VP17.15.4 4.4×10^4 W

Bridging Problem

(a) 14.26 K (b) 0.427 kg/h

? The higher the temperature of a gas, the greater the average kinetic energy of its molecules. How much faster are molecules moving in the air above a frying pan (100°C) than in the surrounding kitchen air (25°C)? (i) 4 times faster; (ii) twice as fast; (iii) 1.25 times as fast; (iv) 1.12 times as fast; (v) 1.06 times as fast.

18 Thermal Properties of Matter

The kitchen is a great place to learn about how the properties of matter depend on temperature. When you boil water in a tea kettle, the increase in temperature produces steam that whistles out of the spout at high pressure. If you forget to poke holes in a potato before baking it, the high-pressure steam produced inside the potato can cause it to explode messily. Water vapor in the air can condense into liquid on the sides of a glass of ice water; if the glass is just out of the freezer, water vapor will solidify and form frost on its sides.

These examples show the relationships among the large-scale or *macroscopic* properties of a substance, such as pressure, volume, temperature, and mass. But we can also describe a substance by using a *microscopic* perspective. This means investigating small-scale quantities such as the masses, speeds, kinetic energies, and momenta of the individual molecules that make up a substance.

The macroscopic and microscopic descriptions are intimately related. For example, the (microscopic) forces that occur when air molecules strike a solid surface (such as your skin) cause (macroscopic) atmospheric pressure. To produce standard atmospheric pressure of 1.01×10^5 Pa, 10^{32} molecules strike your skin every day with an average speed of over 1700 km/h (1000 mi/h)!

In this chapter we'll begin by looking at some macroscopic aspects of matter in general. We'll pay special attention to the *ideal gas,* one of the simplest types of matter to understand. We'll relate the macroscopic properties of an ideal gas to the microscopic behavior of its molecules. We'll also use microscopic ideas to understand the heat capacities of gases and solids. Finally, we'll look at the various phases of matter—gas, liquid, and solid—and the conditions under which each occurs.

LEARNING OUTCOMES

In this chapter, you'll learn...

18.1 How to relate the pressure, volume, and temperature of a gas.
18.2 How the interactions between the molecules of a substance determine the properties of the substance.
18.3 How the pressure and temperature of a gas are related to the kinetic energy of its molecules.
18.4 How the heat capacities of a gas reveal whether its molecules are rotating or vibrating.
18.5 How the speeds of molecules are distributed in a gas.
18.6 What determines whether a substance is a gas, a liquid, or a solid.

You'll need to review...

7.4 Potential energy and force.
11.4 Bulk stress.
12.2 Fluids in equilibrium.
13.3 Escape speed.
14.4 Interatomic forces and oscillations.
17.1–17.6 Temperature, heat, thermal expansion, specific heat, molar heat capacity, phase changes.

18.1 EQUATIONS OF STATE

Quantities such as pressure, volume, temperature, and amount of substance describe the conditions, or *state,* in which a particular material exists. (For example, a tank of medical oxygen has a pressure gauge and a label stating the volume within the tank. We can add a thermometer and put the tank on a scale to measure the mass of oxygen.) These quantities are called **state variables.**

The volume V of a substance is usually determined by its pressure p, temperature T, and amount of substance, described by the mass m_{total} or number of moles n. (We are calling the total mass of a substance m_{total} because later in the chapter we'll use m for the mass of one molecule.) Ordinarily, we can't change one of these variables without causing a change in another. When the tank of oxygen gets hotter, the pressure increases. If the tank gets too hot, it explodes.

In a few cases the relationship among p, V, T, and m_{total} (or n) is simple enough that we can express it as an equation called the **equation of state.** When it's too complicated for that, we can use graphs or numerical tables. Even then, the relationship among the variables still exists; we call it an equation of state even when we don't know the actual equation.

Here's a simple (though approximate) equation of state for a solid material. The temperature coefficient of volume expansion β (see Section 17.4) is the fractional volume change $\Delta V/V_0$ per unit temperature change, and the compressibility k (see Section 11.4) is the negative of the fractional volume change $\Delta V/V_0$ per unit pressure change. If a certain amount of material has volume V_0 when the pressure is p_0 and the temperature is T_0, the volume V at slightly differing pressure p and temperature T is approximately

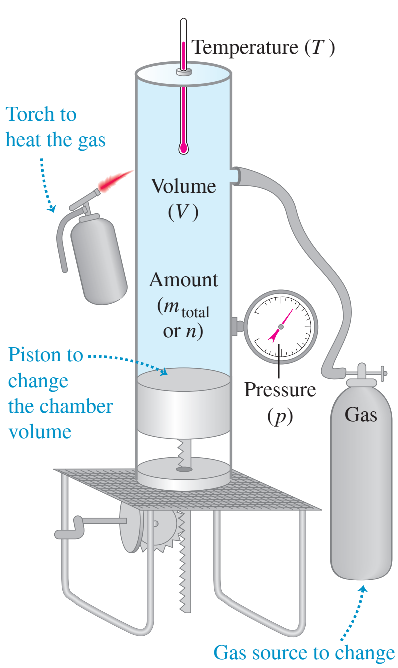
$$V = V_0[1 + \beta(T - T_0) - k(p - p_0)] \tag{18.1}$$

(There is a negative sign in front of the term $k(p - p_0)$ because an *increase* in pressure causes a *decrease* in volume.)

The Ideal-Gas Equation

Another simple equation of state is the one for an *ideal gas*. **Figure 18.1** shows an experimental setup to study the behavior of a gas. The cylinder has a movable piston to vary the volume, the temperature can be varied by heating, and we can pump in any desired amount of gas. We then measure the pressure, volume, temperature, and amount of gas. Note that *pressure* refers both to the force per unit area exerted by the cylinder on the gas and to that exerted by the gas on the cylinder; by Newton's third law, these must be equal.

Figure 18.1 A hypothetical setup for studying the behavior of gases. By heating the gas, varying the volume with a movable piston, and adding more gas, we can control the gas pressure p, volume V, temperature T, and number of moles n.



It is usually easiest to describe the amount of gas in terms of the number of moles n, rather than the mass. (We did this when we defined molar heat capacity in Section 17.5.) The **molar mass** M of a compound (sometimes confusingly called *molecular weight*) is the mass per mole:

$$\underbrace{m_{\text{total}}}_{\text{Total mass of substance}} = \underbrace{n}_{\text{Number of moles of substance}}\ \underbrace{M}_{\text{Molar mass of substance}} \tag{18.2}$$

Hence if we know the number of moles of gas in the cylinder, we can determine the mass of gas from Eq. (18.2).

Measurements of the behavior of various gases lead to three conclusions:

1. The volume V is proportional to the number of moles n. If we double n, keeping pressure and temperature constant, the volume doubles.
2. The volume varies *inversely* with the absolute pressure p. If we double p while holding the temperature T and number of moles n constant, the gas compresses to one-half of its initial volume. In other words, $pV =$ constant when n and T are constant.
3. The pressure is proportional to the *absolute* temperature T. If we double T, keeping the volume and number of moles constant, the pressure doubles. In other words, $p = (\text{constant}) \times T$ when n and V are constant.

We can combine these three relationships into a single **ideal-gas equation:**

Ideal-gas equation:

$$pV = nRT \tag{18.3}$$

Gas pressure: p; Gas volume: V; Number of moles of gas: n; Gas constant: R; Absolute temperature of gas: T

An **ideal gas** is one for which Eq. (18.3) holds precisely for *all* pressures and temperatures. This is an idealized model; it works best at very low pressures and high temperatures, when the gas molecules are far apart and in rapid motion. It is valid within a few percent at moderate pressures (such as a few atmospheres) and at temperatures well above those at which the gas liquefies (**Fig. 18.2**).

Figure **18.2** The ideal-gas equation $pV = nRT$ gives a good description of the air inside an inflated vehicle tire, where the pressure is about 3 atmospheres and the temperature is much too high for nitrogen or oxygen to liquefy. As the tire warms (T increases), the volume V changes only slightly but the pressure p increases.

CAUTION **Use absolute pressure and absolute temperature in the ideal-gas equation** The pressure p in Eq. (18.3), the ideal-gas equation, is the *absolute* pressure, not the gauge pressure (see Section 12.2). Furthermore, the temperature T in Eq. (18.3) is the *absolute* (Kelvin) temperature, not the Celsius temperature. If you use gauge pressure or Celsius temperature in Eq. (18.3), you'll get nonsensical results, such as the pressure of a gas being negative below 0°C = 273.15 K!

We might expect that the proportionality constant R in the ideal-gas equation would have different values for different gases, but it turns out to have the same value for *all* gases, at least at sufficiently high temperature and low pressure. It is called the **gas constant** (or *ideal-gas constant*). In SI units, in which the unit of p is Pa $(1\ \text{Pa} = 1\ \text{N/m}^2)$ and the unit of V is m^3, the numerical value of R (to nine significant figures) is

$$R = 8.31446262\ \text{J/mol} \cdot \text{K}$$

or $R = 8.314\ \text{J/mol} \cdot \text{K}$ to four significant figures. Note that the units of pressure times volume are the same as the units of work or energy (for example, N/m^2 times m^3); that's why R has units of energy per mole per unit of absolute temperature. In chemical calculations, volumes are often expressed in liters (L) and pressures in atmospheres (atm). In this system, to four significant figures,

$$R = 0.08206\ \frac{\text{L} \cdot \text{atm}}{\text{mol} \cdot \text{K}}$$

We can express the ideal-gas equation, Eq. (18.3), in terms of the mass m_{total} of gas, using $m_{\text{total}} = nM$ from Eq. (18.2):

$$pV = \frac{m_{\text{total}}}{M} RT \tag{18.4}$$

From this we can get an expression for the density $\rho = m_{\text{total}}/V$ of the gas:

$$\rho = \frac{pM}{RT} \tag{18.5}$$

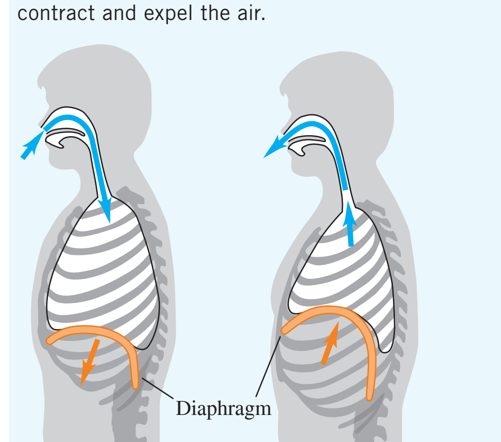
CAUTION **Density vs. pressure** When using Eq. (18.5), be certain that you distinguish between the Greek letter ρ (rho) for density and the letter p for pressure.

For a *constant mass* (or constant number of moles) of an ideal gas the product nR is constant, so the quantity pV/T is also constant. If the subscripts 1 and 2 refer to any two states of the same mass of a gas, then

$$\frac{p_1 V_1}{T_1} = \frac{p_2 V_2}{T_2} = \text{constant} \quad \text{(ideal gas, constant mass)} \tag{18.6}$$

Notice that you don't need the value of R to use this equation.

BIO APPLICATION Respiration and the Ideal-Gas Equation To breathe, you rely on the ideal-gas equation $pV = nRT$. Contraction of the dome-shaped diaphragm muscle increases the volume V of the thoracic cavity (which encloses the lungs), decreasing its pressure p. The lowered pressure causes the lungs to expand and fill with air. (The temperature T is kept constant.) When you exhale, the diaphragm relaxes, allowing the lungs to contract and expel the air.



Inhalation: Diaphragm contracts; lungs expand.

Exhalation: Diaphragm relaxes; lungs contract.

We used the proportionality of pressure to absolute temperature in Chapter 17 to define a temperature scale in terms of pressure in a constant-volume gas thermometer. That may make it seem that the pressure–temperature relationship in the ideal-gas equation, Eq. (18.3), is just a result of the way we define temperature. But the ideal-gas equation also tells us what happens when we change the volume or the amount of substance. Also, we'll see in Chapter 20 that the gas-thermometer scale corresponds closely to a temperature scale that does *not* depend on the properties of any particular material. For now, consider Eq. (18.6) as being based on this genuinely material-independent temperature scale.

PROBLEM-SOLVING STRATEGY 18.1 Ideal Gases

IDENTIFY *the relevant concepts:* Unless the problem states otherwise, you can use the ideal-gas equation to find quantities related to the state of a gas, such as pressure p, volume V, temperature T, and/or number of moles n.

SET UP *the problem* using the following steps:

1. List the known and unknown quantities. Identify the target variables.
2. If the problem concerns only one state of the system, use Eq. (18.3), $pV = nRT$ (or Eq. (18.5), $\rho = pM/RT$ if the problem involves the density ρ rather than n and V).
3. In problems that concern two states (call them 1 and 2) of the same amount of gas, if all but one of the six quantities p_1, p_2, V_1, V_2, T_1, and T_2 are known, use Eq. (18.6), $p_1V_1/T_1 = p_2V_2/T_2 = \text{constant}$. Otherwise, use Eq. (18.3) or Eq. (18.5).

EXECUTE *the solution* as follows:

1. Use consistent units. (SI units are entirely consistent.) The problem statement may make one system of units more convenient than others. Make appropriate unit conversions, such as from atmospheres to pascals or from liters to cubic meters.
2. You may have to convert between mass m_{total} and number of moles n, using $m_{total} = Mn$, where M is the molar mass. If you use Eq. (18.4), you *must* use the same mass units for m_{total} and M. So if M is in grams per mole (the usual units for molar mass), then m_{total} must also be in grams. To use m_{total} in kilograms, you must convert M to kg/mol. For example, the molar mass of oxygen is 32 g/mol or 32×10^{-3} kg/mol.
3. Remember that in the ideal-gas equations, T is always an *absolute* (Kelvin) temperature and p is always an absolute (not gauge) pressure.
4. Solve for the target variables.

EVALUATE *your answer:* Do your results make physical sense? Use benchmarks, such as the result of Example 18.1 below that a mole of an ideal gas at 1 atmosphere pressure occupies a volume of 22.4 liters.

EXAMPLE 18.1 Volume of an ideal gas at STP

WITH VARIATION PROBLEMS

What is the volume of a container that holds exactly 1 mole of an ideal gas at *standard temperature and pressure* (STP), defined as $T = 0°\text{C} = 273.15$ K and $p = 1\text{ atm} = 1.013 \times 10^5$ Pa?

IDENTIFY and SET UP This problem involves the properties of a single state of an ideal gas, so we use Eq. (18.3). We are given the pressure p, temperature T, and number of moles n; our target variable is the corresponding volume V.

EXECUTE From Eq. (18.3), using R in J/mol·K, we get

$$V = \frac{nRT}{p} = \frac{(1\text{ mol})(8.314\text{ J/mol}\cdot\text{K})(273.15\text{ K})}{1.013 \times 10^5\text{ Pa}} = 0.0224\text{ m}^3 = 22.4\text{ L}$$

EVALUATE At STP, 1 mole of an ideal gas occupies 22.4 L. This is the volume of a cube 0.282 m (11.1 in.) on a side, or of a sphere 0.350 m (13.8 in.) in diameter.

KEYCONCEPT The ideal-gas equation, Eq. (18.3), relates the pressure, volume, absolute temperature, and number of moles for a quantity of an ideal gas.

EXAMPLE 18.2 Compressing gas in an automobile engine

WITH VARIATION PROBLEMS

In an automobile engine, a mixture of air and vaporized gasoline is compressed in the cylinders before being ignited. A typical engine has a compression ratio of 9.00 to 1; that is, the gas in the cylinders is compressed to $\frac{1}{9.00}$ of its original volume (**Fig. 18.3**). The intake and exhaust valves are closed during the compression, so the quantity of gas is constant. What is the final temperature of the compressed gas if its initial temperature is 27°C and the initial and final pressures are 1.00 atm and 21.7 atm, respectively?

IDENTIFY and SET UP We must compare two states of the same quantity of ideal gas, so we use Eq. (18.6). In the uncompressed state 1, $p_1 = 1.00$ atm and $T_1 = 27°\text{C} = 300$ K. In the compressed state 2, $p_2 = 21.7$ atm. The cylinder volumes are not given, but we have $V_1 = 9.00V_2$. The temperature T_2 of the compressed gas is the target variable.

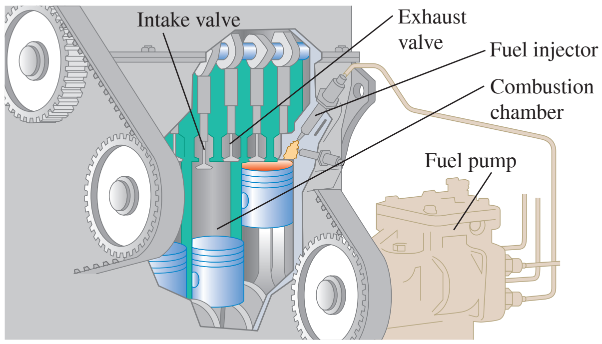


Figure **18.3** Cutaway of an automobile engine. While the air–gasoline mixture is being compressed prior to ignition, both the intake and exhaust valves are in the closed (up) position.

EXECUTE We solve Eq. (18.6) for T_2:

$$T_2 = T_1 \frac{p_2 V_2}{p_1 V_1} = (300\text{ K})\frac{(21.7\text{ atm})V_2}{(1.00\text{ atm})(9.00V_2)} = 723\text{ K} = 450°\text{C}$$

EVALUATE This is the temperature of the air–gasoline mixture *before* the mixture is ignited; when burning starts, the temperature becomes higher still.

KEYCONCEPT For a fixed quantity of an ideal gas, the pressure p, volume V, and absolute temperature T may all change, but the quantity pV/T remains constant.

EXAMPLE 18.3 Mass of air in a scuba tank

WITH VARIATION PROBLEMS

An "empty" aluminum scuba tank contains 11.0 L of air at 21°C and 1 atm. When the tank is filled rapidly from a compressor, the air temperature is 42°C and the gauge pressure is 2.10×10^7 Pa. What mass of air was added? (Air is about 78% nitrogen, 21% oxygen, and 1% miscellaneous; its average molar mass is 28.8 g/mol $= 28.8 \times 10^{-3}$ kg/mol.)

IDENTIFY and SET UP Our target variable is the difference $m_2 - m_1$ between the masses present at the end (state 2) and at the beginning (state 1). We are given the molar mass M of air, so we can use Eq. (18.2) to find the target variable if we know the number of moles present in states 1 and 2. We determine n_1 and n_2 by applying Eq. (18.3) to each state individually.

EXECUTE We convert temperatures to the Kelvin scale by adding 273 and convert the pressure to absolute by adding 1.013×10^5 Pa. The tank's volume is hardly affected by the increased temperature and pressure, so $V_2 = V_1$. From Eq. (18.3), the numbers of moles in the empty tank (n_1) and the full tank (n_2) are as follows:

$$n_1 = \frac{p_1 V_1}{RT_1} = \frac{(1.013 \times 10^5\text{ Pa})(11.0 \times 10^{-3}\text{ m}^3)}{(8.314\text{ J/mol}\cdot\text{K})(294\text{ K})} = 0.46\text{ mol}$$

$$n_2 = \frac{p_2 V_2}{RT_2} = \frac{(2.11 \times 10^7\text{ Pa})(11.0 \times 10^{-3}\text{ m}^3)}{(8.314\text{ J/mol}\cdot\text{K})(315\text{ K})} = 88.6\text{ mol}$$

We added $n_2 - n_1 = 88.6\text{ mol} - 0.46\text{ mol} = 88.1$ mol to the tank. From Eq. (18.2), the added mass is $M(n_2 - n_1) = (28.8 \times 10^{-3}\text{ kg/mol})(88.1\text{ mol}) = 2.54$ kg.

EVALUATE The added mass is not insubstantial: You could certainly use a scale to determine whether the tank was empty or full.

KEYCONCEPT You can determine the mass of a quantity of an ideal gas from its pressure, volume, and absolute temperature, and the molar mass of the gas.

EXAMPLE 18.4 Variation of atmospheric pressure with elevation

WITH VARIATION PROBLEMS

Find the variation of atmospheric pressure with elevation in the earth's atmosphere. Assume that at all elevations, $T = 0°$C and $g = 9.80\text{ m/s}^2$.

IDENTIFY and SET UP As the elevation y increases, both the atmospheric pressure p and the density ρ decrease. Hence we have *two* unknown functions of y; to solve for them, we need two independent equations. One is the ideal-gas equation, Eq. (18.5), which is expressed in terms of p and ρ. The other is Eq. (12.4), the relationship that we found in Section 12.2 among p, ρ, and y in a fluid in equilibrium: $dp/dy = -\rho g$. We are told to assume that g and T are the same at all elevations; we also assume that the atmosphere has the same chemical composition, and hence the same molar mass M, at all heights. We combine the two equations and solve for $p(y)$.

EXECUTE We substitute $\rho = pM/RT$ into $dp/dy = -\rho g$, separate variables, and integrate, letting p_1 be the pressure at elevation y_1 and p_2 be the pressure at y_2:

$$\frac{dp}{dy} = -\frac{pM}{RT}g$$

$$\int_{p_1}^{p_2}\frac{dp}{p} = -\frac{Mg}{RT}\int_{y_1}^{y_2} dy$$

$$\ln\frac{p_2}{p_1} = -\frac{Mg}{RT}(y_2 - y_1)$$

$$\frac{p_2}{p_1} = e^{-Mg(y_2 - y_1)/RT}$$

Now let $y_1 = 0$ be at sea level and let the pressure at that point be $p_0 = 1.013 \times 10^5$ Pa. Then the pressure p at any height y is

$$p = p_0 e^{-Mgy/RT}$$

Continued

EVALUATE According to our calculation, the pressure decreases exponentially with elevation. The graph in **Fig. 18.4** shows that the slope dp/dy becomes less negative with greater elevation. That result makes sense, since $dp/dy = -\rho g$ and the density also decreases with elevation. At the summit of Mount Everest, where $y = 8848$ m,

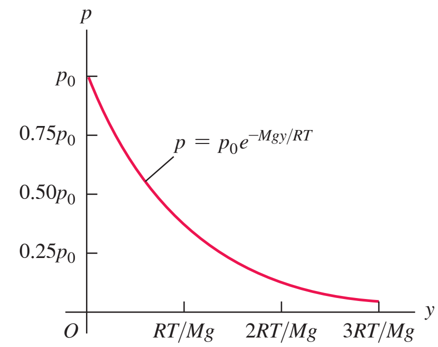
$$\frac{Mgy}{RT} = \frac{(28.8 \times 10^{-3}\ \text{kg/mol})(9.80\ \text{m/s}^2)(8848\ \text{m})}{(8.314\ \text{J/mol}\cdot\text{K})(273\ \text{K})} = 1.10$$

$$p = (1.013 \times 10^5\ \text{Pa})e^{-1.10} = 0.337 \times 10^5\ \text{Pa} = 0.33\ \text{atm}$$

The assumption of constant temperature isn't realistic, and g decreases a little with increasing elevation (see Challenge Problem 18.84). Even so, this example shows why most mountaineers carry oxygen on Mount Everest. It also shows why jet airliners, which typically fly at altitudes of 8000 to 12,000 m, *must* have pressurized cabins for passenger comfort and health.

KEYCONCEPT The ideal-gas equation can also be expressed as a relationship among the pressure, density, molar mass, and absolute temperature of an ideal gas [Eq. (12.4)].

Figure 18.4 The variation of atmospheric pressure p with elevation y, assuming a constant temperature T.



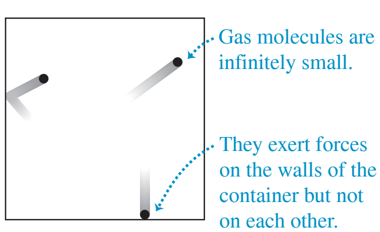
The van der Waals Equation

In Section 18.3 we'll obtain the ideal-gas equation, Eq. (18.3), from a simple molecular model that ignores the volumes of the molecules themselves and the attractive forces between them (**Fig. 18.5a**). Another equation of state, the **van der Waals equation,** makes approximate corrections for these two omissions (Fig. 18.5b). This equation was developed by the 19th-century Dutch physicist J. D. van der Waals; the interaction between atoms that we discussed in Section 14.4 is named the *van der Waals interaction*. The van der Waals equation is

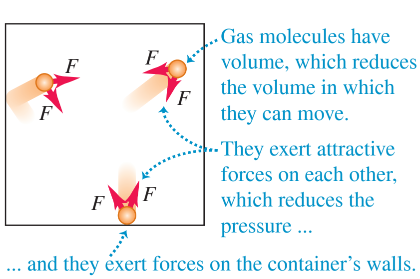
$$\left(p + \frac{an^2}{V^2}\right)(V - nb) = nRT \qquad (18.7)$$

Figure 18.5 A gas as modeled by (a) the ideal-gas equation and (b) the van der Waals equation.

(a) An idealized model of a gas



(b) A more realistic model of a gas



The constants a and b are different for different gases. Roughly speaking, b represents the volume of a mole of molecules; the total volume of the molecules is nb, and the volume remaining in which the molecules can move is $V - nb$. The constant a depends on the attractive intermolecular forces, which reduce the pressure of the gas by *pulling* the molecules together as they *push* on the walls of the container. The decrease in pressure is proportional to the number of molecules per unit volume in a layer near the wall (which are exerting the pressure on the wall) and is also proportional to the number per unit volume in the next layer beyond the wall (which are doing the attracting). Hence the decrease in pressure due to intermolecular forces is proportional to n^2/V^2.

When n/V is small (that is, when the gas is *dilute*), the average distance between molecules is large, the corrections in the van der Waals equation become insignificant, and Eq. (18.7) reduces to the ideal-gas equation. As an example, for carbon dioxide gas (CO_2) the constants in the van der Waals equation are $a = 0.364\ \text{J}\cdot\text{m}^3/\text{mol}^2$ and $b = 4.27 \times 10^{-5}\ \text{m}^3/\text{mol}$. We saw in Example 18.1 that 1 mole of an ideal gas at $T = 0°\text{C} = 273.15$ K and $p = 1\ \text{atm} = 1.013 \times 10^5$ Pa occupies a volume $V = 0.0224\ \text{m}^3$; according to Eq. (18.7), 1 mole of CO_2 occupying this volume at this temperature would be at a pressure 532 Pa less than 1 atm, a difference of only 0.5% from the ideal-gas value.

pV-Diagrams

We could in principle represent the *p*-*V*-*T* relationship graphically as a *surface* in a three-dimensional space with coordinates p, V, and T. This representation is useful (see Section 18.6), but ordinary two-dimensional graphs are usually more convenient. One of the most useful of these is a set of graphs of pressure as a function of volume, each for a particular constant temperature. Such a diagram is called a ***pV*-diagram.** Each curve, representing behavior at a specific temperature, is called an **isotherm,** or a *pV-isotherm*.

Figure 18.6 shows pV-isotherms for a constant amount of an ideal gas. Since $p = nRT/V$ from Eq. (18.3), along an isotherm (constant T) the pressure p is inversely proportional to the volume V and the isotherms are hyperbolic curves.

Figure 18.7 shows a pV-diagram for a material that *does not* obey the ideal-gas equation. At temperatures below T_c the isotherms develop flat regions in which we can compress the material (that is, reduce the volume V) without increasing the pressure p. Observation shows that the gas is *condensing* from the vapor (gas) to the liquid phase. The flat parts of the isotherms in the shaded area of Fig. 18.7 represent conditions of liquid–vapor *phase equilibrium.* As the volume decreases, more and more material goes from vapor to liquid, but the pressure does not change. (To keep the temperature constant during condensation, we have to remove the heat of vaporization, discussed in Section 17.6.)

When we compress such a gas at a constant temperature T_2 in Fig. 18.7, it is vapor until point a is reached. Then it begins to liquefy; as the volume decreases further, more material liquefies, and *both* the pressure and the temperature remain constant. At point b, all the material is in the liquid state. After this, any further compression requires a very rapid rise of pressure, because liquids are in general much less compressible than gases. At a lower constant temperature T_1, similar behavior occurs, but the condensation begins at lower pressure and greater volume than at the constant temperature T_2. At temperatures greater than T_c, *no* phase transition occurs as the material is compressed; at the highest temperatures, such as T_4, the curves resemble the ideal-gas curves of Fig. 18.6. We call T_c the *critical temperature* for this material. In Section 18.6 we'll discuss what happens to the phase of the gas above the critical temperature.

We'll use pV-diagrams often in the next two chapters. We'll show that the *area* under a pV-curve (whether or not it is an isotherm) represents the *work* done by the system during a volume change. This work, in turn, is directly related to heat transfer and changes in the *internal energy* of the system.

Figure **18.6** Isotherms, or constant-temperature curves, for a constant amount of an ideal gas. The highest temperature is T_4; the lowest is T_1. This is a graphical representation of the ideal-gas equation of state.

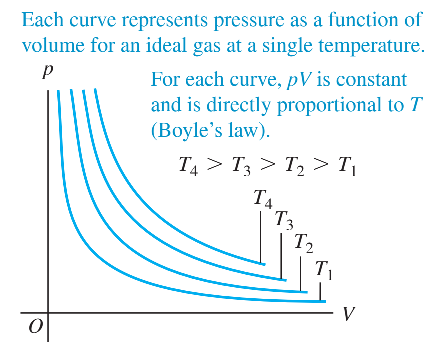
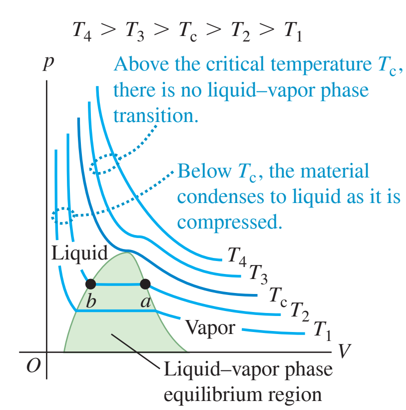


Figure **18.7** A pV-diagram for a nonideal gas, showing isotherms for temperatures above and below the critical temperature T_c. The liquid–vapor equilibrium region is shown as a green shaded area. At still lower temperatures the material might undergo phase transitions from liquid to solid or from gas to solid; these are not shown here.



TEST YOUR UNDERSTANDING OF SECTION 18.1 Rank the following ideal gases in order from highest to lowest number of moles: (i) Pressure $p = 1$ atm, volume $V = 1$ L, temperature $T = 300$ K; (ii) $p = 2$ atm, $V = 1$ L, $T = 300$ K; (iii) $p = 1$ atm, $V = 2$ L, $T = 300$ K; (iv) $p = 1$ atm, $V = 1$ L, $T = 600$ K; (v) $p = 2$ atm, $V = 1$ L, $T = 600$ K.

ANSWER

❙ **(ii) and (iii) (tie), (i) and (v) (tie), (iv)** We can rewrite the ideal-gas equation, Eq. (18.3), as $n = pV/RT$. This tells us that the number of moles n is proportional to the pressure and volume and inversely proportional to the absolute temperature. Hence, compared to (i), the number of moles in each case is (ii) $(2)(1)/(1) = 2$ times as much, (iii) $(1)(2)/(1) = 2$ times as much, (iv) $(1)(1)/(2) = \frac{1}{2}$ as much, and (v) $(2)(1)/(2) = 1$ time as much (that is, equal).

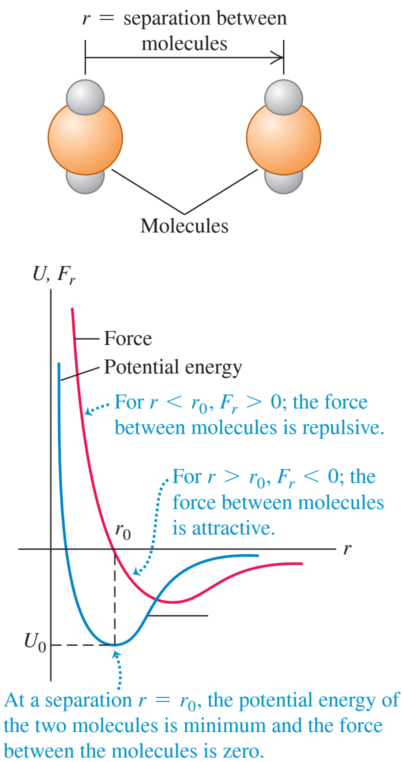
18.2 MOLECULAR PROPERTIES OF MATTER

We have studied several properties of matter in bulk, including elasticity, density, surface tension, heat capacities, and equations of state. Now we want to look in more detail at the relationship of bulk behavior to *molecular* structure. We begin with a general discussion of the molecular structure of matter. Then in the next two sections we develop the kinetic-molecular model of an ideal gas, obtaining from this molecular model the equation of state and an expression for heat capacity.

Molecules and Intermolecular Forces

Any specific chemical compound is made up of identical **molecules.** The smallest molecules contain one atom each and are of the order of 10^{-10} m in size; the largest contain many atoms and are at least 10,000 times larger. In gases the molecules move nearly independently; in liquids and solids they are held together by intermolecular forces. These forces arise from interactions among the electrically charged particles that make up the molecules. Gravitational forces between molecules are negligible in comparison with electric forces.

Figure 18.8 How the force between molecules and their potential energy of interaction depend on their separation r.



The interaction of two *point* electric charges is described by a force (repulsive for like charges, attractive for unlike charges) with a magnitude proportional to $1/r^2$, where r is the distance between the points. We'll study this relationship, called *Coulomb's law,* in Chapter 21. Molecules are *not* point charges but complex structures containing both positive and negative charge, and their interactions are more complex. The force between molecules in a gas varies with the distance r between molecules somewhat as shown in **Fig. 18.8**, where a positive F_r corresponds to a repulsive force and a negative F_r to an attractive force. When molecules are far apart, the intermolecular forces are very small and usually attractive. As a gas is compressed and its molecules are brought closer together, the attractive forces increase. The intermolecular force becomes zero at an equilibrium spacing r_0, corresponding roughly to the spacing between molecules in the liquid and solid states. In liquids and solids, relatively large pressures are needed to compress the substance appreciably. This shows that at molecular distances slightly *less* than r_0, the forces become *repulsive* and relatively large.

Figure 18.8 also shows the potential energy as a function of r. This function has a *minimum* at r_0, where the force is zero. The two curves are related by $F_r(r) = -dU/dr$, as we showed in Section 7.4. Such a potential-energy function is often called a **potential well.** A molecule at rest at a distance r_0 from a second molecule would need an additional energy $|U_0|$, the "depth" of the potential well, to "escape" to an indefinitely large value of r.

Molecules are always in motion; their kinetic energies usually increase with temperature. At very low temperatures the average kinetic energy of a molecule may be much *less* than the depth of the potential well. The molecules then condense into the liquid or solid phase with average intermolecular spacings of about r_0. But at higher temperatures the average kinetic energy becomes larger than the depth $|U_0|$ of the potential well. Molecules can then escape the intermolecular force and become free to move independently, as in the gaseous phase of matter.

In *solids,* molecules vibrate about more or less fixed points. (See Section 17.4.) In a crystalline solid these points are arranged in a *crystal lattice.* **Figure 18.9** shows the cubic crystal structure of sodium chloride, and **Fig. 18.10** shows a scanning tunneling microscope image of individual silicon atoms on the surface of a crystal.

In a *liquid,* the intermolecular distances are usually only slightly greater than in the solid phase of the same substance, but the molecules have much greater freedom of movement. Liquids show regularity of structure only in the immediate neighborhood of a few molecules.

Figure 18.9 Schematic representation of the cubic crystal structure of sodium chloride (ordinary salt).

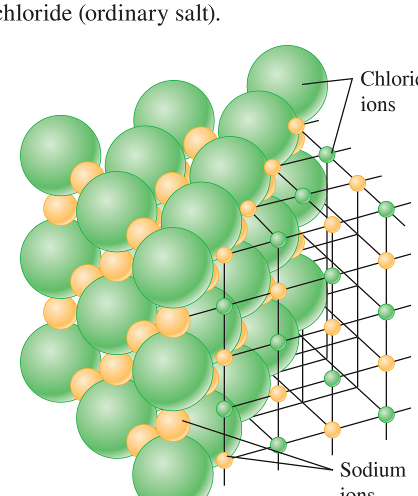
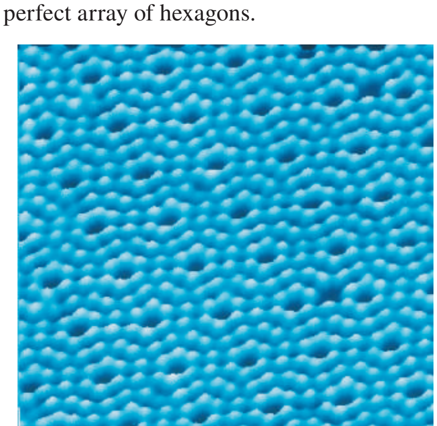


Figure 18.10 A scanning tunneling microscope image of the surface of a silicon crystal. The area shown is only 9.0 nm (9.0×10^{-9} m) across. Each blue "bead" is one silicon atom; these atoms are arranged in a (nearly) perfect array of hexagons.



The molecules of a *gas* are usually widely separated and so have only very small attractive forces. A gas molecule moves in a straight line until it collides with another molecule or with a wall of the container. In molecular terms, an *ideal gas* is a gas whose molecules exert *no* attractive forces on each other (see Fig. 18.5a) and therefore have no *potential* energy.

At low temperatures, most common substances are in the solid phase. As the temperature rises, a substance melts and then vaporizes. From a molecular point of view, these transitions are in the direction of increasing molecular kinetic energy. Thus temperature and molecular kinetic energy are closely related.

Moles and Avogadro's Number

We have used the mole as a measure of quantity of substance. One **mole** of any pure chemical element or compound contains a definite number of molecules, the same number for all elements and compounds. The official SI definition is:

One mole is the amount of substance that contains $6.02214076 \times 10^{23}$ elementary entities.

In our discussion, the "elementary entities" are molecules. (In a monatomic substance such as carbon or helium, each molecule is a single atom.) Until 2018 the number of elementary entities per mole was defined in terms of the number of atoms in 0.012 kilogram of carbon-12. The new definition is based on the redefinition of the kilogram (see Section 1.3) and on careful measurements of nearly perfect spheres of pure silicon-28.

The number of molecules in a mole is called **Avogadro's number,** denoted by N_A. The numerical value of N_A is *defined* to be

$$N_A = 6.02214076 \times 10^{23}\ \text{molecules/mol} \quad \text{(Avogadro's number)}$$

The *molar mass* M of a compound is the mass of 1 mole. It is equal to the mass m of a single molecule multiplied by Avogadro's number:

$$\underbrace{M}_{\text{Molar mass of a substance}} = \underbrace{N_A}_{\text{Avogadro's number}}\, \underbrace{m}_{\text{Mass of a molecule of substance}} \qquad (18.8)$$

When the molecule consists of a single atom, the term *atomic mass* is often used instead of molar mass.

EXAMPLE 18.5 Atomic and molecular mass

Find the mass of a single hydrogen atom and of a single oxygen molecule.

IDENTIFY and SET UP This problem involves the relationship between the mass of a molecule or atom (our target variable) and the corresponding molar mass M. We use Eq. (18.8) in the form $m = M/N_A$ and the values of the atomic masses from the periodic table of the elements (see Appendix D).

EXECUTE For atomic hydrogen the atomic mass (molar mass) is $M_H = 1.008$ g/mol, so the mass m_H of a single hydrogen atom is

$$m_H = \frac{1.008\ \text{g/mol}}{6.022 \times 10^{23}\ \text{atoms/mol}} = 1.674 \times 10^{-24}\ \text{g/atom}$$

For oxygen the atomic mass is 16.0 g/mol, so for the diatomic (two-atom) oxygen molecule the molar mass is 32.0 g/mol. Then the mass of a single oxygen molecule is

$$m_{O_2} = \frac{32.0\ \text{g/mol}}{6.022 \times 10^{23}\ \text{molecules/mol}} = 53.1 \times 10^{-24}\ \text{g/molecule}$$

EVALUATE We note that the values in Appendix D are for the *average* atomic masses of a natural sample of each element. Such a sample may contain several *isotopes* of the element, each with a different atomic mass. Natural samples of hydrogen and oxygen are almost entirely made up of just one isotope.

KEYCONCEPT To find the mass of a single molecule of a substance, divide the molar mass of that substance by Avogadro's number (the number of molecules in a mole).

TEST YOUR UNDERSTANDING OF SECTION 18.2 Suppose you could adjust the value of r_0 for the molecules of a certain chemical compound (Fig. 18.8) by turning a dial. If you doubled the value of r_0, the density of the solid form of this compound would become (i) twice as great; (ii) four times as great; (iii) eight times as great; (iv) $\frac{1}{2}$ as great; (v) $\frac{1}{4}$ as great; (vi) $\frac{1}{8}$ as great.

ANSWER

(vi) The value of r_0 determines the equilibrium separation of the molecules in the solid phase, so doubling r_0 means that the separation doubles as well. Hence a solid cube of this compound might grow from 1 cm on a side to 2 cm on a side. The volume would then be $2^3 = 8$ times larger, and the density (mass divided by volume) would be $\frac{1}{8}$ as great.

18.3 KINETIC-MOLECULAR MODEL OF AN IDEAL GAS

The goal of any molecular theory of matter is to understand the *macroscopic* properties of matter in terms of its atomic or molecular structure and behavior. Once we have this understanding, we can design materials to have specific desired properties. Theories have led to the development of high-strength steels, semiconductor materials for electronic devices, and countless other materials essential to contemporary technology.

Let's consider a simple molecular model of an ideal gas. This *kinetic-molecular model* represents the gas as a large number of particles bouncing around in a closed container. In this section we use the kinetic-molecular model to understand how the ideal-gas equation of state, Eq. (18.3), is related to Newton's laws. In the following section we use the kinetic-molecular model to predict the molar heat capacity of an ideal gas. We'll go on to elaborate the model to include "particles" that are not points but have a finite size.

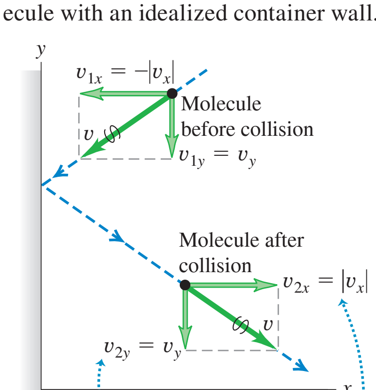
Our discussion of the kinetic-molecular model has several steps, and you may need to go over them several times. Don't get discouraged!

Here are the assumptions of our model:

1. A container with volume V contains a very large number N of identical molecules, each with mass m.
2. The molecules behave as point particles that are small compared to the size of the container and to the average distance between molecules.
3. The molecules are in constant motion. Each molecule collides occasionally with a wall of the container. These collisions are perfectly elastic.
4. The container walls are rigid and infinitely massive and do not move.

CAUTION Molecules vs. moles Don't confuse N, the number of *molecules* in the gas, with n, the number of *moles*. The number of molecules is equal to the number of moles multiplied by Avogadro's number: $N = nN_A$.

Figure 18.11 Elastic collision of a molecule with an idealized container wall.



Collisions and Gas Pressure

During collisions the molecules exert *forces* on the walls of the container; this is the origin of the *pressure* that the gas exerts. In a typical collision (**Fig. 18.11**) the velocity component parallel to the wall is unchanged, and the component perpendicular to the wall reverses direction but does not change in magnitude.

We'll first determine the *number* of collisions that occur per unit time for a certain area A of wall. Then we find the total momentum change associated with these collisions and the force needed to cause this momentum change. From this we can determine the pressure (force per unit area) and compare to the ideal-gas equation. We'll find a direct connection between the temperature of the gas and the kinetic energy of its molecules.

To begin, we'll assume that all molecules in the gas have the same *magnitude* of x-velocity, $|v_x|$. Later we'll see that our results don't depend on making this overly simplistic assumption.

As Fig. 18.11 shows, for each collision the x-component of velocity changes from $-|v_x|$ to $+|v_x|$. So the x-component of momentum p_x changes from $-m|v_x|$ to $+m|v_x|$, and the *change* in p_x is $m|v_x| - (-m|v_x|) = 2m|v_x|$.

If a molecule is going to collide with a given wall area A during a small time interval dt, then at the beginning of dt it must be within a distance $|v_x|\,dt$ from the wall (**Fig. 18.12**) and it must be headed toward the wall. So the number of molecules that collide with A during dt is equal to the number of molecules within a cylinder with base area A and length $|v_x|\,dt$ that have their x-velocity aimed toward the wall. The volume of such a cylinder is $A|v_x|\,dt$. Assuming that the number of molecules per unit volume (N/V) is uniform, the *number* of molecules in this cylinder is $(N/V)(A|v_x|\,dt)$. On the average, half of these molecules are moving toward the wall and half are moving away from it. So the number of collisions with A during dt is

$$\frac{1}{2}\left(\frac{N}{V}\right)(A|v_x|\,dt)$$

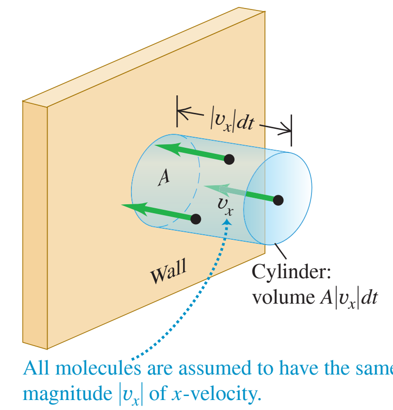


Figure 18.12 For a molecule to strike the wall in area A during a time interval dt, the molecule must be headed for the wall and be within the shaded cylinder of length $|v_x|\,dt$ at the beginning of the interval.

For the system of all molecules in the gas, the total momentum change dP_x during dt is the *number* of collisions multiplied by $2m|v_x|$:

$$dP_x = \frac{1}{2}\left(\frac{N}{V}\right)(A|v_x|\,dt)(2m|v_x|) = \frac{NAmv_x^2\,dt}{V} \tag{18.9}$$

(We are using capital P for total momentum and lowercase p for pressure. Be careful!) We wrote v_x^2 rather than $|v_x|^2$ in the final expression because the square of the absolute value of a number is equal to the square of that number. The *rate* of change of momentum component P_x is

$$\frac{dP_x}{dt} = \frac{NAmv_x^2}{V} \tag{18.10}$$

According to Newton's second law, this rate of change of momentum equals the force exerted by the wall area A on the gas molecules. From Newton's *third* law this is equal and opposite to the force exerted *on* the wall *by* the molecules. Pressure p is the magnitude of the force exerted on the wall per unit area:

$$p = \frac{F}{A} = \frac{Nmv_x^2}{V} \tag{18.11}$$

The pressure exerted by the gas depends on the number of molecules per volume (N/V), the mass m per molecule, and the speed of the molecules.

Pressure and Molecular Kinetic Energies

We mentioned that $|v_x|$ is really *not* the same for all the molecules. But we could have sorted the molecules into groups having the same $|v_x|$ within each group, then added up the resulting contributions to the pressure. The net effect of all this is just to replace v_x^2 in Eq. (18.11) by the *average* value of v_x^2, which we denote by $(v_x^2)_{av}$. We can relate $(v_x^2)_{av}$ to the *speeds* of the molecules. The speed v of a molecule is related to the velocity components v_x, v_y, and v_z by

$$v^2 = v_x^2 + v_y^2 + v_z^2$$

We can average this relationship over all molecules:

$$(v^2)_{av} = (v_x^2)_{av} + (v_y^2)_{av} + (v_z^2)_{av}$$

But there is no real difference in our model between the x-, y-, and z-directions. (Molecular speeds are very fast in a typical gas, so the effects of gravity are negligibly small.) It follows that $(v_x^2)_{av}$, $(v_y^2)_{av}$, and $(v_z^2)_{av}$ must all be *equal*. Hence $(v^2)_{av}$ is equal to $3(v_x^2)_{av}$ and

$$(v_x^2)_{av} = \tfrac{1}{3}(v^2)_{av}$$

so Eq. (18.11) becomes

$$pV = \tfrac{1}{3}Nm(v^2)_{av} = \tfrac{2}{3}N\left[\tfrac{1}{2}m(v^2)_{av}\right] \tag{18.12}$$

Figure **18.13** Summer air (top) is warmer than winter air (bottom); that is, the average translational kinetic energy of air molecules is greater in summer.

CAUTION **The Boltzmann constant, Avogadro's number, the gas constant, and the Kelvin scale** In our discussion we began with the gas constant R and Avogadro's number N_A and used them to define the Boltzmann constant k. In fact, in the International System (SI; see Section 1.3), both N_A and k are fundamental quantities with defined values. The value of the gas constant is the product of these two defined values: $R = N_A k$. Furthermore, the value of k provides the official definition of the Kelvin scale: If the temperature increases by one kelvin, the quantity kT increases by exactly 1.380649×10^{-23} J.

We notice that $\frac{1}{2}m(v^2)_{av}$ is the average translational kinetic energy of a single molecule. The product of this and the total number of molecules N equals the total random kinetic energy K_{tr} of translational motion of all the molecules. (The notation K_{tr} reminds us that this is the energy of *translational* motion. There may also be energies associated with molecular rotation and vibration.) The product pV equals two-thirds of the total translational kinetic energy:

$$pV = \tfrac{2}{3}K_{tr} \tag{18.13}$$

Now compare Eq. (18.13) to the ideal-gas equation $pV = nRT$, Eq. (18.3), which is based on experimental studies of gas behavior. For the two equations to agree, we must have

Average translational kinetic energy of an ideal gas

$$K_{tr} = \tfrac{3}{2}nRT \tag{18.14}$$

Number of moles of gas
Absolute temperature of gas
Gas constant

So K_{tr} is *directly proportional* to the absolute temperature T (**Fig. 18.13**).

The average translational kinetic energy of a single molecule is the total translational kinetic energy K_{tr} of all molecules divided by the number of molecules, N:

$$\frac{K_{tr}}{N} = \tfrac{1}{2}m(v^2)_{av} = \frac{3nRT}{2N}$$

Also, the total number of molecules N is the number of moles n multiplied by Avogadro's number N_A, so $N = nN_A$ and $n/N = 1/N_A$. Thus the above equation becomes

$$\frac{K_{tr}}{N} = \tfrac{1}{2}m(v^2)_{av} = \tfrac{3}{2}\left(\frac{R}{N_A}\right)T \tag{18.15}$$

The ratio R/N_A is called the **Boltzmann constant,** k:

$$k = \frac{R}{N_A} = \frac{8.314\ \text{J/mol}\cdot\text{K}}{6.022 \times 10^{23}\ \text{molecules/mol}} = 1.381 \times 10^{-23}\ \text{J/molecule}\cdot\text{K}$$

(The exact numerical value of k is *defined* to be 1.380649×10^{-23} J/molecule · K.) In terms of k we can rewrite Eq. (18.15) as

Average translational kinetic energy of a gas molecule:

$$\tfrac{1}{2}m(v^2)_{av} = \tfrac{3}{2}kT \tag{18.16}$$

Mass of a molecule
Absolute temperature of gas
Average value of the square of molecular speeds
Boltzmann constant

This shows that the average translational kinetic energy *per molecule* depends only on the temperature, not on the pressure, volume, or kind of molecule. We can obtain the average translational kinetic energy *per mole* by multiplying Eq. (18.16) by Avogadro's number and using the relationship $M = N_A m$:

$$N_A\tfrac{1}{2}m(v^2)_{av} = \tfrac{1}{2}M(v^2)_{av} = \tfrac{3}{2}RT \quad \text{(average translational kinetic energy per mole of gas)} \tag{18.17}$$

The translational kinetic energy of a mole of an ideal gas depends only on T.

Finally, it can be helpful to rewrite the ideal-gas equation on a "per-molecule" basis. We use $N = N_A n$ and $R = N_A k$ to obtain this alternative form:

$$pV = NkT \tag{18.18}$$

This shows that we can think of the Boltzmann constant k as a gas constant on a "per-molecule" basis instead of the usual "per-mole" basis for R.

Molecular Speeds

From Eqs. (18.16) and (18.17) we can obtain expressions for the square root of $(v^2)_{av}$, called the **root-mean-square speed** (or **rms speed**) v_{rms}:

$$\underbrace{v_{rms}}_{\text{Root-mean-square speed of a gas molecule}} = \sqrt{\underbrace{(v^2)_{av}}_{\text{Average value of the square of molecular speeds}}} = \sqrt{\frac{3\overbrace{k}^{\text{Boltzmann constant}}\overbrace{T}^{\text{Absolute temperature of gas}}}{\underbrace{m}_{\text{Mass of a molecule}}}} = \sqrt{\frac{3\underbrace{R}_{\text{Gas constant}}T}{\underbrace{M}_{\text{Molar mass}}}} \tag{18.19}$$

It might seem more natural to give the *average* speed rather than v_{rms}, but v_{rms} follows more directly from Eqs. (18.16) and (18.17). To compute the rms speed, we square each molecular speed, add, divide by the number of molecules, and take the square root; v_{rms} is the *root* of the *mean* of the *squares*.

Equations (18.16) and (18.19) show that at a given temperature T, gas molecules of different mass m have the same average kinetic energy but different root-mean-square speeds. On average, the nitrogen molecules ($M = 28$ g/mol) in the air around you are moving faster than are the oxygen molecules ($M = 32$ g/mol). Hydrogen molecules ($M = 2$ g/mol) are fastest of all; this is why there is hardly any hydrogen in the earth's atmosphere, despite its being the most common element in the universe (**Fig. 18.14**). A sizable fraction of any H_2 molecules in the atmosphere would have speeds greater than the earth's escape speed of 1.12×10^4 m/s (calculated in Example 13.5 in Section 13.3) and would escape into space. The heavier, slower-moving gases cannot escape so easily, which is why they predominate in our atmosphere.

The assumption that individual molecules undergo perfectly elastic collisions with the container wall is a little too simple. In most cases, molecules actually adhere to the wall for a short time and then leave again with speeds that are characteristic of the temperature *of the wall*. However, the gas and the wall are ordinarily in thermal equilibrium and have the same temperature. So there is no net energy transfer between gas and wall, and our conclusions remain valid.

Figure **18.14** While hydrogen is a desirable fuel for vehicles, it is only a trace constituent of our atmosphere (0.00005% by volume). Hence hydrogen fuel has to be generated by electrolysis of water, which is itself an energy-intensive process.

PROBLEM-SOLVING STRATEGY 18.2 Kinetic-Molecular Theory

IDENTIFY *the relevant concepts:* Use the results of the kinetic-molecular model to relate the macroscopic properties of a gas, such as temperature and pressure, to microscopic properties, such as molecular speeds.

SET UP *the problem* using the following steps:

1. List knowns and unknowns; identify the target variables.
2. Choose appropriate equation(s) from among Eqs. (18.14), (18.16), and (18.19).

EXECUTE *the solution* as follows: Maintain consistency in units.

1. The usual units for molar mass M are grams per mole; these units are often omitted in tables. In equations such as Eq. (18.19), when you use SI units you must express M in kilograms per mole. For example, for oxygen $M_{O_2} = 32$ g/mol $= 32 \times 10^{-3}$ kg/mol.
2. Are you working on a "per-molecule" basis (with m, N, and k) or a "per-mole" basis (with M, n, and R)? To check units, think of N as having units of "molecules"; then m has units of mass per molecule, and k has units of joules per molecule per kelvin. Similarly, n has units of moles; then M has units of mass per mole and R has units of joules per mole per kelvin.
3. Remember that T is always *absolute* (Kelvin) temperature.

EVALUATE *your answer:* Are your answers reasonable? Here's a benchmark: Typical molecular speeds at room temperature are several hundred meters per second.

EXAMPLE 18.6 Molecular kinetic energy and v_{rms}

WITH VARIATION PROBLEMS

(a) What is the average translational kinetic energy of an ideal-gas molecule at 27°C? (b) What is the total random translational kinetic energy of the molecules in 1 mole of this gas? (c) What is the rms speed of oxygen molecules at this temperature?

IDENTIFY and SET UP This problem involves the translational kinetic energy of an ideal gas on a per-molecule and per-mole basis, as well as the root-mean-square molecular speed v_{rms}. We are given $T = 27°C = 300$ K and $n = 1$ mol; we use the molecular mass m for oxygen. We use Eq. (18.16) to determine the average kinetic energy of a molecule, Eq. (18.14) to find the total molecular kinetic energy K_{tr} of 1 mole, and Eq. (18.19) to find v_{rms}.

EXECUTE (a) From Eq. (18.16),

$$\tfrac{1}{2}m(v^2)_{av} = \tfrac{3}{2}kT = \tfrac{3}{2}(1.38 \times 10^{-23}\ \text{J/K})(300\ \text{K})$$
$$= 6.21 \times 10^{-21}\ \text{J}$$

(b) From Eq. (18.14), the kinetic energy of one mole is

$$K_{tr} = \tfrac{3}{2}nRT = \tfrac{3}{2}(1\ \text{mol})(8.314\ \text{J/mol}\cdot\text{K})(300\ \text{K}) = 3740\ \text{J}$$

(c) We found the mass per molecule m and molar mass M of molecular oxygen in Example 18.5. Using Eq. (18.19), we can calculate v_{rms} in two ways:

$$v_{rms} = \sqrt{\frac{3kT}{m}} = \sqrt{\frac{3(1.38 \times 10^{-23}\ \text{J/K})(300\ \text{K})}{5.31 \times 10^{-26}\ \text{kg}}}$$
$$= 484\ \text{m/s} = 1740\ \text{km/h} = 1080\ \text{mi/h}$$

$$v_{rms} = \sqrt{\frac{3RT}{M}} = \sqrt{\frac{3(8.314\ \text{J/mol}\cdot\text{K})(300\ \text{K})}{32.0 \times 10^{-3}\ \text{kg/mol}}} = 484\ \text{m/s}$$

EVALUATE The answer in part (a) does not depend on the mass of the molecule. We can check our result in part (b) by noting that the translational kinetic energy per mole must be equal to the product of the average translational kinetic energy per molecule from part (a) and Avogadro's number N_A: $K_{tr} = (6.21 \times 10^{-21}\ \text{J/molecule}) \times (6.022 \times 10^{23}\ \text{molecules}) = 3740$ J.

KEYCONCEPT The average translational kinetic energy of a molecule in an ideal gas at absolute temperature T is $\tfrac{3}{2}kT$, where k is the Boltzmann constant, no matter what the mass of the molecules. However, the root-mean-square speed v_{rms} of molecules (that is, the square root of the average value of v^2) in an ideal gas does depend on the mass per molecule m: $v_{rms} = \sqrt{\dfrac{3kT}{m}}$.

EXAMPLE 18.7 Calculating rms and average speeds

WITH VARIATION PROBLEMS

Five gas molecules have speeds of 500, 600, 700, 800, and 900 m/s. What is the rms speed? What is the *average* speed?

IDENTIFY and SET UP We use the definitions of the root mean square and the average of a collection of numbers. To find v_{rms}, we square each speed, find the average (mean) of the squares, and take the square root of the result. We find v_{av} as usual.

EXECUTE The average value of v^2 and the resulting v_{rms} for the five molecules are

$$(v^2)_{av} = \frac{500^2 + 600^2 + 700^2 + 800^2 + 900^2}{5}\ \text{m}^2/\text{s}^2$$
$$= 5.10 \times 10^5\ \text{m}^2/\text{s}^2$$
$$v_{rms} = \sqrt{(v^2)_{av}} = 714\ \text{m/s}$$

The average speed v_{av} is

$$v_{av} = \frac{500 + 600 + 700 + 800 + 900}{5}\ \text{m/s} = 700\ \text{m/s}$$

EVALUATE In general v_{rms} and v_{av} are *not* the same. Roughly speaking, v_{rms} gives greater weight to the higher speeds than does v_{av}.

KEYCONCEPT In general, for any collection of numbers, the root-mean-square value is *not* equal to the average value. For molecules in an ideal gas, the root-mean-square speed is always greater than the average speed.

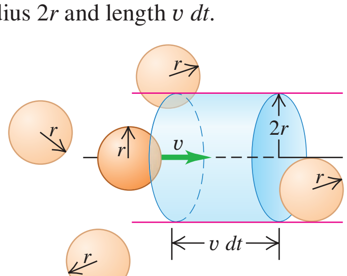


Figure 18.15 In a time dt a molecule with radius r will collide with any other molecule within a cylindrical volume of radius $2r$ and length $v\,dt$.

Collisions Between Molecules

We have ignored the possibility that two gas molecules might collide. If they are really points, they *never* collide. But consider a more realistic model in which the molecules are rigid spheres with radius r. How often do they collide with other molecules? How far do they travel, on average, between collisions? We can get approximate answers from the following rather primitive model.

Consider N spherical molecules with radius r in a volume V. Suppose only one molecule is moving. When it collides with another molecule, the distance between centers is $2r$. Suppose we draw a cylinder with radius $2r$, with its axis parallel to the velocity of the molecule (**Fig. 18.15**). The moving molecule collides with any other molecule whose center is inside this cylinder. In a short time dt a molecule with speed v travels a distance $v\,dt$;

during this time it collides with any molecule that is in the cylindrical volume of radius $2r$ and length $v\,dt$. The volume of the cylinder is $4\pi r^2 v\,dt$. There are N/V molecules per unit volume, so the number dN with centers in this cylinder is

$$dN = 4\pi r^2 v\,dt\,N/V$$

Thus the number of collisions *per unit time* is

$$\frac{dN}{dt} = \frac{4\pi r^2 v N}{V}$$

This result assumes that only one molecule is moving. It turns out that collisions are more frequent when all the molecules move at once, and the above equation has to be multiplied by a factor of $\sqrt{2}$:

$$\frac{dN}{dt} = \frac{4\pi\sqrt{2}\,r^2 v N}{V}$$

The average time t_{mean} between collisions, called the *mean free time,* is the reciprocal of this expression:

$$t_{\text{mean}} = \frac{V}{4\pi\sqrt{2}\,r^2 v N} \qquad (18.20)$$

The average distance traveled between collisions is called the **mean free path.** In our model, this is just the molecule's speed v multiplied by t_{mean}:

$$\lambda = v t_{\text{mean}} = \frac{V}{4\pi\sqrt{2}\,r^2 N} \qquad (18.21)$$

Mean free path of a gas molecule
Speed of molecule
Volume of gas
Mean free time between collisions
Radius of a molecule
Number of molecules in gas

The mean free path λ (the Greek letter lambda) is inversely proportional to the number of molecules per unit volume (N/V) and inversely proportional to the cross-sectional area πr^2 of a molecule; the more molecules there are and the larger the molecule, the shorter the mean distance between collisions (**Fig. 18.16**). Note that the mean free path *does not* depend on the speed of the molecule.

We can express Eq. (18.21) in terms of macroscopic properties of the gas, using the ideal-gas equation in the form of Eq. (18.18), $pV = NkT$. We find

$$\lambda = \frac{kT}{4\pi\sqrt{2}\,r^2 p} \qquad (18.22)$$

If the temperature is increased at constant pressure, the gas expands, the average distance between molecules increases, and λ increases. If the pressure is increased at constant temperature, the gas compresses and λ decreases.

Figure **18.16** If you try to walk through a crowd, your mean free path—the distance you can travel on average without running into another person—depends on how large the people are and how closely they are spaced.

EXAMPLE 18.8 Calculating mean free path

WITH VARIATION PROBLEMS

(a) Estimate the mean free path of a molecule of air at 27°C and 1 atm. Model the molecules as spheres with radius $r = 2.0 \times 10^{-10}$ m. (b) Estimate the mean free time of an oxygen molecule with $v = v_{\text{rms}}$ at 27°C and 1 atm.

IDENTIFY and SET UP This problem uses the concepts of mean free path and mean free time (our target variables). We use Eq. (18.22) to determine the mean free path λ. We then use $\lambda = v t_{\text{mean}}$ in Eq. (18.21), with $v = v_{\text{rms}}$, to find the mean free time t_{mean}.

EXECUTE (a) From Eq. (18.22),

$$\lambda = \frac{kT}{4\pi\sqrt{2}r^2 p} = \frac{(1.38 \times 10^{-23}\ \text{J/K})(300\ \text{K})}{4\pi\sqrt{2}(2.0 \times 10^{-10}\ \text{m})^2(1.01 \times 10^5\ \text{Pa})}$$
$$= 5.8 \times 10^{-8}\ \text{m}$$

Continued

(b) From Example 18.6, for oxygen at 27°C the root-mean-square speed is $v_{rms} = 484\ \text{m/s}$, so the mean free time for a molecule with this speed is

$$t_{mean} = \frac{\lambda}{v} = \frac{5.8 \times 10^{-8}\ \text{m}}{484\ \text{m/s}} = 1.2 \times 10^{-10}\ \text{s}$$

This molecule undergoes about 10^{10} collisions per second!

EVALUATE Note that from Eqs. (18.21) and (18.22) the mean free *path* doesn't depend on the molecule's speed, but the mean free *time* does. Slower molecules have a longer average time interval t_{mean} between collisions than do fast ones, but the average *distance* λ between collisions is the same no matter what the molecule's speed. Our answer to part (a) says that the molecule doesn't go far between collisions, but the mean free path is still several hundred times the molecular radius r.

KEYCONCEPT The mean free path and mean free time are, respectively, the average distance and average time that a molecule in a gas travels between collisions with other molecules. Both depend on the temperature and pressure of the gas and the radius of a molecule. The mean free time also depends on the speed of the molecule; the mean free path does not.

TEST YOUR UNDERSTANDING OF SECTION 18.3 Rank the following gases in order from (a) highest to lowest rms speed of molecules and (b) highest to lowest average translational kinetic energy of a molecule: (i) oxygen ($M = 32.0$ g/mol) at 300 K; (ii) nitrogen ($M = 28.0$ g/mol) at 300 K; (iii) oxygen at 330 K; (iv) nitrogen at 330 K.

ANSWER

(a) (iv), (ii), (iii), (i); (b) (iii) and (iv) (tie), (i) and (ii) (tie) (a) Equation (18.19) tells us that $v_{rms} = \sqrt{3RT/M}$, so the rms speed is proportional to the square root of the ratio of absolute temperature T to molar mass M. Compared to (i) oxygen at 300 K, v_{rms} in the other cases is (ii) $\sqrt{(32.0\ \text{g/mol})/(28.0\ \text{g/mol})} = 1.07$ times faster, (iii) $\sqrt{(330\ \text{K})/(300\ \text{K})} = 1.05$ times faster, and (iv) $\sqrt{(330\ \text{K})(32.0\ \text{g/mol})/(300\ \text{K})(28.0\ \text{g/mol})} = 1.12$ times faster. (b) From Eq. (18.16), the average translational kinetic energy per molecule is $\frac{1}{2}m(v^2)_{av} = \frac{3}{2}kT$, which is directly proportional to T and independent of M. We have $T = 300$ K for cases (i) and (ii) and $T = 330$ K for cases (iii) and (iv), so $\frac{1}{2}m(v^2)_{av}$ has equal values for cases (iii) and (iv) and equal (but smaller) values for cases (i) and (ii).

18.4 HEAT CAPACITIES

When we introduced the concept of heat capacity in Section 17.5, we talked about ways to *measure* the specific heat or molar heat capacity of a particular material. Now we'll see how to *predict* these on theoretical grounds.

Heat Capacities of Gases

The basis of our analysis is that heat is *energy* in transit. When we add heat to a substance, we are increasing its molecular energy. We'll assume that the volume of the gas remains constant; if we were to let the gas expand, it would do work by pushing on the moving walls of its container, and this additional energy transfer would have to be included in our calculations. We'll return to this more general case in Chapter 19. For now we are concerned with C_V, the molar heat capacity *at constant volume.*

In the simple kinetic-molecular model of Section 18.3 the molecular energy consists of only the translational kinetic energy K_{tr} of the pointlike molecules. This energy is directly proportional to the absolute temperature T, as shown by Eq. (18.14), $K_{tr} = \frac{3}{2}nRT$. When the temperature changes by a small amount dT, the corresponding change in kinetic energy is

$$dK_{tr} = \tfrac{3}{2}nR\,dT \qquad (18.23)$$

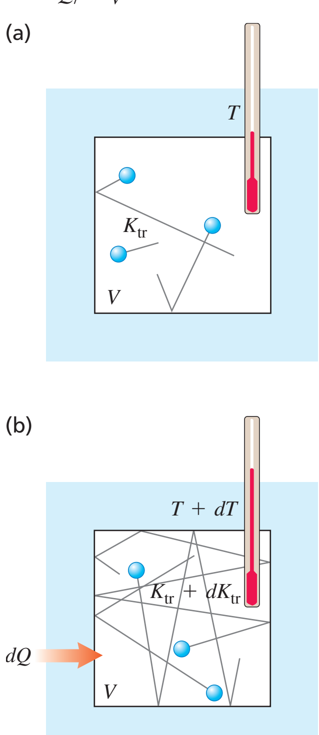
From the definition of molar heat capacity at constant volume, C_V (see Section 17.5), we also have

$$dQ = nC_V\,dT \qquad (18.24)$$

where dQ is the heat input needed for a temperature change dT. Now if K_{tr} represents the total molecular energy, as we have assumed, then dQ and dK_{tr} must be *equal* (**Fig. 18.17**). From Eqs. (18.23) and (18.24), this says

$$nC_V\,dT = \tfrac{3}{2}nR\,dT$$

Figure 18.17 (a) A fixed volume V of a monatomic ideal gas. (b) When an amount of heat dQ is added to the gas, the total translational kinetic energy increases by $dK_{tr} = dQ$ and the temperature increases by $dT = dQ/nC_V$.



Molar heat capacity at constant volume, ideal gas of point particles
$$C_V = \frac{3}{2}R \quad \text{(Gas constant)} \tag{18.25}$$

This surprisingly simple result says that the molar heat capacity at constant volume is $\frac{3}{2}R$ for *any* gas whose molecules can be represented as points.

Does Eq. (18.25) agree with experiment? In SI units, Eq. (18.25) gives

$$C_V = \tfrac{3}{2}(8.314\ \text{J/mol}\cdot\text{K}) = 12.47\ \text{J/mol}\cdot\text{K}$$

For comparison, **Table 18.1** gives measured values of C_V for several gases. We see that for *monatomic* gases our prediction is right on the money, but that it is way off for diatomic and polyatomic gases.

This comparison tells us that our point-molecule model is good enough for monatomic gases but that for diatomic and polyatomic molecules we need something more sophisticated. For example, we can picture a diatomic molecule as *two* point masses, like a little elastic dumbbell (see **Fig. 18.18**), with an interaction force between the atoms of the kind shown in Fig. 18.8. Such a molecule can have additional kinetic energy associated with *rotation* about axes through its center of mass. The atoms may also vibrate along the line joining them, with additional kinetic and potential energies.

When heat flows into a *monatomic* gas at constant volume, *all* of the added energy goes into an increase in random *translational* molecular kinetic energy. Equation (18.23) shows that this gives rise to an increase in temperature. But when the temperature is increased by the same amount in a *diatomic* or *polyatomic* gas, additional heat is needed to supply the increased rotational and vibrational energies. Thus polyatomic gases have *larger* molar heat capacities than monatomic gases, as Table 18.1 shows.

But how do we know how much energy is associated with each additional kind of motion of a complex molecule, compared to the translational kinetic energy? The new principle that we need is called the principle of **equipartition of energy.** It can be derived from sophisticated statistical-mechanics considerations; that derivation is beyond our scope, and we'll treat the principle as an axiom.

The principle of equipartition of energy states that each velocity component (either linear or angular) has, on average, an associated kinetic energy per molecule of $\frac{1}{2}kT$, or one-half the product of the Boltzmann constant and the absolute temperature. The number of velocity components needed to describe the motion of a molecule completely is called the number of **degrees of freedom.** For a monatomic gas, there are three degrees of freedom (for the velocity components v_x, v_y, and v_z); this gives a total average kinetic energy per molecule of $3\left(\frac{1}{2}kT\right)$, consistent with Eq. (18.16).

For a *diatomic* molecule there are two possible axes of rotation, perpendicular to each other and to the molecule's axis. (We don't include rotation about the molecule's own axis because in ordinary collisions there is no way for this rotational motion to change.) If we add two rotational degrees of freedom for a diatomic molecule, the average total kinetic energy per molecule is $\frac{5}{2}kT$ instead of $\frac{3}{2}kT$. The total kinetic energy of n moles is $K_{\text{total}} = nN_A\left(\frac{5}{2}kT\right) = \frac{5}{2}n(kN_A)T = \frac{5}{2}nRT$, and the molar heat capacity (at constant volume) is

Molar heat capacity at constant volume, ideal diatomic gas
$$C_V = \frac{5}{2}R \quad \text{(Gas constant)} \tag{18.26}$$

In SI units,

$$C_V = \tfrac{5}{2}(8.314\ \text{J/mol}\cdot\text{K}) = 20.79\ \text{J/mol}\cdot\text{K}$$

This value is close to the measured values for diatomic gases in Table 18.1.

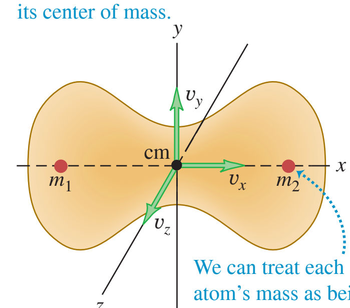
Vibrational motion can also contribute to the heat capacities of gases. Molecular bonds can stretch and bend, and the resulting vibrations lead to additional degrees of freedom and additional energies. For most diatomic gases, however, vibration does *not* contribute appreciably to heat capacity. The reason for this involves some concepts of quantum mechanics. Briefly, vibrational energy can change only in finite steps. If the energy change of the first step is much larger than the energy possessed by most molecules, then nearly all

TABLE 18.1 Molar Heat Capacities of Gases

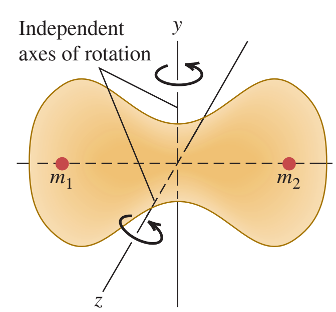
Type of Gas	Gas	C_V (J/mol · K)
Monatomic	He	12.47
	Ar	12.47
Diatomic	H_2	20.42
	N_2	20.76
	O_2	20.85
	CO	20.85
Polyatomic	CO_2	28.46
	SO_2	31.39
	H_2S	25.95

Figure **18.18** Motions of a diatomic molecule.

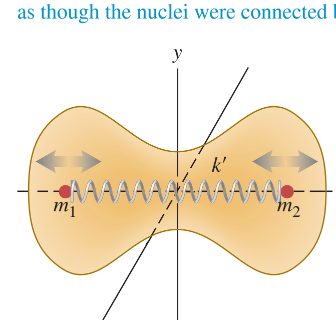
(a) **Translational motion.** The molecule moves as a whole; its velocity may be described as the x-, y-, and z-velocity components of its center of mass.



(b) **Rotational motion.** The molecule rotates about its center of mass. This molecule has two independent axes of rotation.



(c) **Vibrational motion.** The molecule oscillates as though the nuclei were connected by a spring.



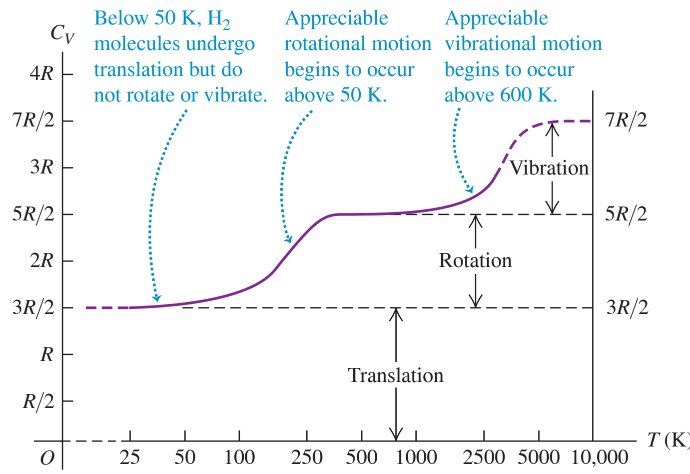


Figure 18.19 Experimental values of C_V, the molar heat capacity at constant volume, for hydrogen gas (H_2). The temperature is plotted on a logarithmic scale.

the molecules remain in the minimum-energy state of motion. Changing the temperature does not change their average vibrational energy appreciably, and the vibrational degrees of freedom are said to be "frozen out." In more complex molecules the gaps between permitted energy levels can be much smaller, and then vibration *does* contribute to heat capacity. The rotational energy of a molecule also changes by finite steps, but they are usually much smaller; the "freezing out" of rotational degrees of freedom occurs only in rare instances.

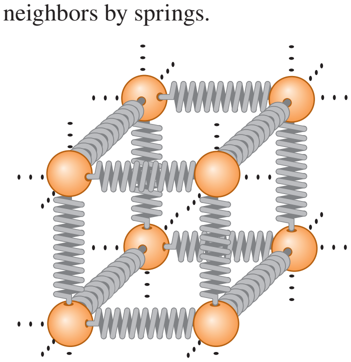
In Table 18.1 the large values of C_V for polyatomic molecules show the effects of vibrational energy. In addition, a molecule with three or more atoms that are not in a straight line has *three* rotational degrees of freedom.

From this discussion we expect heat capacities to be temperature-dependent, generally increasing with increasing temperature. **Figure 18.19** is a graph of the temperature dependence of C_V for hydrogen gas (H_2), showing the temperatures at which the rotational and vibrational energies begin to contribute.

Heat Capacities of Solids

We can carry out a similar heat-capacity analysis for a crystalline solid. Consider a crystal consisting of N identical atoms (a *monatomic solid*). Each atom is bound to an equilibrium position by interatomic forces. Solid materials are elastic, so forces must permit stretching and bending of the bonds. We can think of a crystal as an array of atoms connected by little springs (**Fig. 18.20**).

Figure 18.20 To visualize the forces between neighboring atoms in a crystal, envision every atom as being attached to its neighbors by springs.



Each atom can *vibrate* around its equilibrium position and has three degrees of freedom, corresponding to its three components of velocity. According to the equipartition principle, each atom has an average kinetic energy of $\frac{1}{2}kT$ for each degree of freedom. In addition, there is *potential* energy associated with the elastic deformation. For a simple harmonic oscillator (discussed in Chapter 14) it is not hard to show that the average kinetic energy is *equal* to the average potential energy. In our model of a crystal, each atom is a three-dimensional harmonic oscillator; it can be shown that the equality of average kinetic and potential energies also holds here, provided that the "spring" forces obey Hooke's law.

Thus we expect each atom to have an average kinetic energy $\frac{3}{2}kT$ and an average potential energy $\frac{3}{2}kT$, or an average total energy $3kT$ per atom. If the crystal contains N atoms or n moles, its total energy is

$$E_{\text{total}} = 3NkT = 3nRT \tag{18.27}$$

From this we conclude that the molar heat capacity of a crystal should be

Molar heat capacity of an ideal monatomic solid (rule of Dulong and Petit)
$$C_V = 3R \tag{18.28}$$
Gas constant

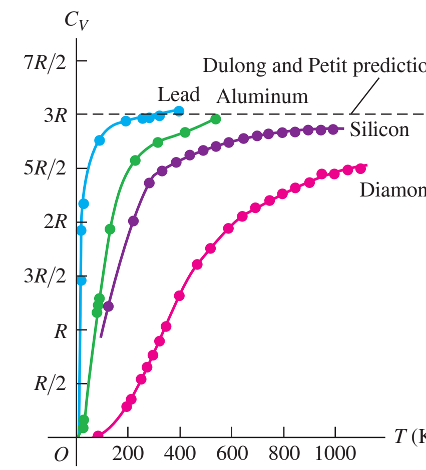
In SI units,

$$C_V = (3)(8.314 \text{ J/mol} \cdot \text{K}) = 24.9 \text{ J/mol} \cdot \text{K}$$

We have *derived* the **rule of Dulong and Petit,** which we encountered as an *empirical* finding in Section 17.5: Monatomic solids all have molar heat capacities of about 25 J/mol · K. The agreement is only approximate, but given the very simple nature of our model, it is quite significant.

At low temperatures, the heat capacities of most solids *decrease* with decreasing temperature (**Fig. 18.21**) for the same reason that vibrational degrees of freedom of molecules are frozen out at low temperatures. At very low temperatures the quantity kT is much *smaller* than the smallest energy step the vibrating atoms can take. Hence most of the atoms remain in their lowest energy states because the next higher energy level is out of reach. The average vibrational energy per atom is then *less* than $3kT$, and the heat capacity per molecule is *less* than $3k$. At higher temperatures when kT is *large* in comparison to the minimum energy step, the equipartition principle holds, and the total heat capacity is $3k$ per molecule or $3R$ per mole as the rule of Dulong and Petit predicts. Quantitative understanding of the temperature variation of heat capacities was one of the triumphs of quantum mechanics during its initial development in the 1920s.

Figure **18.21** Experimental values of C_V for lead, aluminum, silicon, and diamond. At high temperatures, C_V for each solid approaches about $3R$, in agreement with the rule of Dulong and Petit. At low temperatures, C_V is much less than $3R$.



TEST YOUR UNDERSTANDING OF SECTION 18.4 A cylinder with a fixed volume contains hydrogen gas (H_2) at 25 K. You then add heat to the gas at a constant rate until its temperature reaches 500 K. Does the temperature of the gas increase at a constant rate? Why or why not? If not, does the temperature increase most rapidly near the beginning or near the end of this process?

ANSWER

no; near the beginning Adding a small amount of heat dQ to the gas changes the temperature by dT, where $dQ = nC_V dT$ from Eq. (18.24). Figure 18.19 shows that C_V for H_2 varies with temperature between 25 K and 500 K, so a given amount of heat gives rise to different amounts of temperature change during the process. Hence the temperature will *not* increase at a constant rate. The temperature change $dT = dQ/nC_V$ is inversely proportional to C_V, so the temperature increases most rapidly at the beginning of the process when the temperature is lowest and C_V is smallest (see Fig. 18.19).

18.5 MOLECULAR SPEEDS

As we mentioned in Section 18.3, the molecules in a gas don't all have the same speed. **Figure 18.22** shows one experimental scheme for measuring the distribution of molecular speeds. A substance is vaporized in a hot oven; molecules of the vapor escape through an aperture in the oven wall and into a vacuum chamber. A series of slits blocks all molecules except those in a narrow beam, which is aimed at a pair of rotating disks. A molecule passing through the slit in the first disk is blocked by the second disk unless it arrives just as the slit in the second disk is lined up with the beam. The disks function as a speed selector that passes only molecules within a certain narrow speed range. This range can be varied by changing the disk rotation speed, and we can measure how many molecules lie within each of various speed ranges.

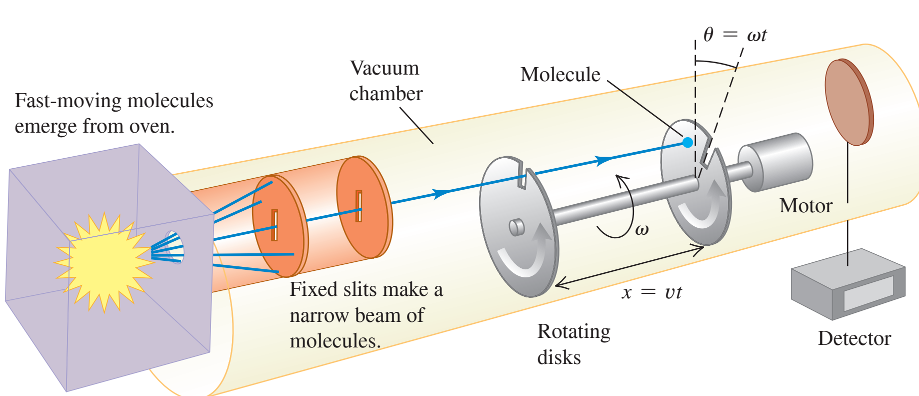


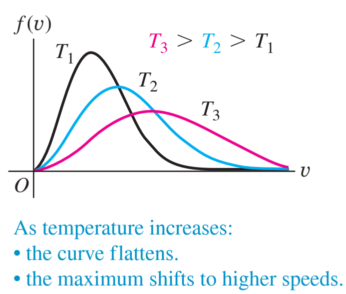
Figure **18.22** A molecule with a speed v passes through the slit in the first rotating disk. When the molecule reaches the second rotating disk, the disks have rotated through the offset angle θ. If $v = \omega x/\theta$, the molecule passes through the slit in the second rotating disk and reaches the detector.

To describe the results of such measurements, we define a function $f(v)$ called a *distribution function*. If we observe a total of N molecules, the number dN having speeds in the range between v and $v + dv$ is given by

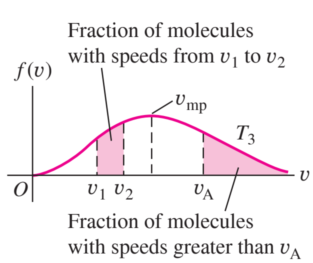
$$dN = Nf(v)\,dv \tag{18.29}$$

Figure 18.23 (a) Curves of the Maxwell–Boltzmann distribution function $f(v)$ for three temperatures. (b) The shaded areas under the curve represent the fractions of molecules within certain speed ranges. The most probable speed v_{mp} for a given temperature is at the peak of the curve.

(a)



(b)



The *probability* that a randomly chosen molecule will have a speed in the interval v to $v + dv$ is $f(v)dv$. Hence $f(v)$ is the probability per unit speed *interval;* it is *not* the probability that a molecule has speed exactly equal to v. Since a probability is a pure number, $f(v)$ has units of reciprocal speed (s/m).

Figure 18.23a shows distribution functions for three different temperatures. At each temperature the height of the curve for any value of v is proportional to the number of molecules with speeds near v. The peak of the curve represents the *most probable speed* v_{mp} for the corresponding temperature. As the temperature increases, the average molecular kinetic energy increases, and so the peak of $f(v)$ shifts to higher and higher speeds.

Figure 18.23b shows that the area under a curve between any two values of v represents the fraction of all the molecules having speeds in that range. Every molecule must have *some* value of v, so the integral of $f(v)$ over all v must be unity for any T.

If we know $f(v)$, we can calculate the most probable speed v_{mp}, the average speed v_{av}, and the rms speed v_{rms}. To find v_{mp}, we simply find the point where $df/dv = 0$; this gives the value of the speed where the curve has its peak. To find v_{av}, we take the number $Nf(v)\,dv$ having speeds in each interval dv, multiply each number by the corresponding speed v, add all these products (by integrating over all v from zero to infinity), and finally divide by N. That is,

$$v_{av} = \int_0^\infty vf(v)\,dv \tag{18.30}$$

We can find the rms speed in a similar way; the average of v^2 is

$$(v^2)_{av} = \int_0^\infty v^2 f(v)\,dv \tag{18.31}$$

and v_{rms} is the square root of this.

The Maxwell–Boltzmann Distribution

The function $f(v)$ describing the actual distribution of molecular speeds is called the **Maxwell–Boltzmann distribution.** It can be derived from statistical-mechanics considerations, but that derivation is beyond our scope. Here is the result:

Mass of a gas molecule
Molecular speed
Mass of a gas molecule
Molecular speed

Maxwell–Boltzmann distribution function

$$f(v) = 4\pi\left(\frac{m}{2\pi kT}\right)^{3/2} v^2 e^{-mv^2/2kT} \tag{18.32}$$

Boltzmann constant

Boltzmann constant
Absolute temperature of gas

We can also express this function in terms of the translational kinetic energy of a molecule, which we denote by ϵ; that is, $\epsilon = \frac{1}{2}mv^2$. We invite you to verify that when this is substituted into Eq. (18.32), the result is

$$f(\epsilon) = \frac{8\pi}{m}\left(\frac{m}{2\pi kT}\right)^{3/2} \epsilon e^{-\epsilon/kT} \tag{18.33}$$

This form shows that the exponent in the Maxwell–Boltzmann distribution function is $-\epsilon/kT$; the shape of the curve is determined by the relative magnitude of ϵ and kT at any point. You can prove that the *peak* of each curve occurs where $\epsilon = kT$, corresponding to a most probable speed v_{mp}:

$$v_{mp} = \sqrt{\frac{2kT}{m}} \tag{18.34}$$

To find the average speed, we substitute Eq. (18.32) into Eq. (18.30), make a change of variable $v^2 = x$, and integrate by parts. The result is

$$v_{\text{av}} = \sqrt{\frac{8kT}{\pi m}} \tag{18.35}$$

Finally, to find the rms speed, we substitute Eq. (18.32) into Eq. (18.31). Evaluating the resulting integral takes some mathematical acrobatics, but we can find it in a table of integrals. The result is

$$v_{\text{rms}} = \sqrt{\frac{3kT}{m}} \tag{18.36}$$

This result agrees with Eq. (18.19); it *must* agree if the Maxwell–Boltzmann distribution is to be consistent with our kinetic-theory calculations.

Table 18.2 shows the fraction of all the molecules in an ideal gas that have speeds *less than* various multiples of v_{rms}. These numbers were obtained by numerical integration; they are the same for all ideal gases.

TABLE 18.2 Fractions of Molecules in an Ideal Gas with Speeds Less Than Various Multiples of v_{rms}

v/v_{rms}	Fraction
0.20	0.011
0.40	0.077
0.60	0.218
0.80	0.411
1.00	0.608
1.20	0.771
1.40	0.882
1.60	0.947
1.80	0.979
2.00	0.993

The distribution of molecular speeds in liquids is similar, although not identical, to that for gases. We can understand evaporation and the vapor pressure of a liquid on this basis. Suppose a molecule must have a speed at least as great as v_A in Fig. 18.23b to escape from the surface of a liquid into the adjacent vapor. The number of such molecules, represented by the area under the "tail" of each curve (to the right of v_A), increases rapidly with temperature. Thus the rate at which molecules can escape is strongly temperature-dependent. This process is balanced by one in which molecules in the vapor phase collide inelastically with the surface and are trapped into the liquid phase. The number of molecules suffering this fate per unit time is proportional to the pressure in the vapor phase. Phase equilibrium between liquid and vapor occurs when these two competing processes proceed at the same rate. So if the molecular speed distributions are known for various temperatures, we can make a theoretical prediction of vapor pressure as a function of temperature. When liquid evaporates, it's the high-speed molecules that escape from the surface. The ones that are left have less energy on average; this gives us a molecular view of evaporative cooling.

Rates of chemical reactions are often strongly temperature-dependent, and the reason is contained in the Maxwell–Boltzmann distribution. When two reacting molecules collide, the reaction can occur only when the molecules are close enough for their electrons to interact strongly. This requires a minimum energy, called the *activation energy,* and thus a minimum molecular speed. Figure 18.23a shows that the number of molecules in the high-speed tail of the curve increases rapidly with temperature. Thus we expect the rate of any reaction with an activation energy to increase rapidly with temperature.

BIO APPLICATION Activation Energy and Moth Activity This hawkmoth of genus *Manduca* cannot fly if the temperature of its muscles is below 29°C. The reason is that the enzyme-catalyzed reactions that power aerobic metabolism and enable muscle action require a minimum molecular energy (activation energy). Just like the molecules in an ideal gas, at low temperatures very few of the molecules involved in these reactions have high energy. As the temperature increases, more molecules have the required minimum energy and the reactions take place at a greater rate. Above 29°C, enough power is generated to allow the hawkmoth to fly.

TEST YOUR UNDERSTANDING OF SECTION 18.5 A quantity of gas containing N molecules has a speed distribution function $f(v)$. How many molecules have speeds between v_1 and $v_2 > v_1$? (i) $\int_0^{v_2} f(v)\,dv - \int_0^{v_1} f(v)\,dv$; (ii) $N\left[\int_0^{v_2} f(v)\,dv - \int_0^{v_1} f(v)\,dv\right]$; (iii) $\int_0^{v_1} f(v)\,dv - \int_0^{v_2} f(v)\,dv$; (iv) $N\left[\int_0^{v_1} f(v)\,dv - \int_0^{v_2} f(v)\,dv\right]$; (v) none of these.

ANSWER

(ii) Figure 18.23b shows that the *fraction* of molecules with speeds between v_1 and v_2 equals the area under the curve of $f(v)$ versus v from $v = v_1$ to $v = v_2$. This is equal to the integral $\int_{v_1}^{v_2} f(v)\,dv$, which in turn is equal to the difference between the integrals $\int_0^{v_2} f(v)\,dv$ (the fraction of molecules with speeds from 0 to v_2) and $\int_0^{v_1} f(v)\,dv$ (the fraction of molecules with speeds from 0 to the slower speed v_1). The *number* of molecules with speeds from v_1 to v_2 equals the fraction of molecules in this speed range multiplied by N, the total number of molecules.

18.6 PHASES OF MATTER

An ideal gas is the simplest system to analyze from a molecular viewpoint because we ignore the interactions between molecules. But those interactions are the very thing that makes matter condense into the liquid and solid phases under some conditions. So it's not surprising that theoretical analysis of liquid and solid structure and behavior is a lot more

Figure **18.24** A typical pT phase diagram, showing regions of temperature and pressure at which the various phases exist and where phase changes occur.

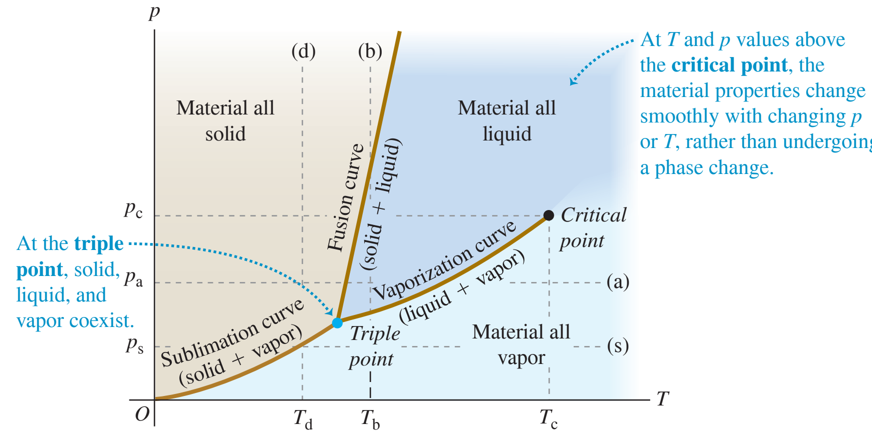


Figure **18.25** Atmospheric pressure on earth is higher than the triple-point pressure of water (see line (a) in Fig. 18.24). Depending on the temperature, water can exist as a vapor (in the atmosphere), as a liquid (in the ocean), or as a solid (like the iceberg shown here).

complicated than that for gases. We won't try to go far here with a microscopic picture, but we can talk in general about phases of matter, phase equilibrium, and phase transitions.

In Section 17.6 we learned that each phase is stable in only certain ranges of temperature and pressure. A transition from one phase to another ordinarily requires **phase equilibrium** between the two phases, and for a given pressure this occurs at only one specific temperature. We can represent these conditions on a graph with axes p and T, called a **phase diagram; Fig. 18.24** shows an example. Each point on the diagram represents a pair of values of p and T.

Only a single phase can exist at each point in Fig. 18.24, except for points on the solid lines, where two phases can coexist in phase equilibrium. The fusion curve separates the solid and liquid areas and represents possible conditions of solid-liquid phase equilibrium. The vaporization curve separates the liquid and vapor areas, and the sublimation curve separates the solid and vapor areas. All three curves meet at the **triple point,** the only condition under which all three phases can coexist (**Fig. 18.25**). In Section 17.3 we used the triple-point temperature of water to define the Kelvin temperature scale. **Table 18.3** gives triple-point data for several substances.

If we heat a substance at a constant pressure p_a, it goes through a series of states represented by the horizontal line (a) in Fig. 18.24. The melting and boiling temperatures at this pressure are the temperatures at which the line intersects the fusion and vaporization curves, respectively. When the pressure is p_s, constant-pressure heating transforms a substance from solid directly to vapor. This process is called *sublimation;* the intersection of line (s) with the sublimation curve gives the temperature T_s at which it occurs for a pressure p_s. At any pressure less than the triple-point pressure, no liquid phase is possible. The triple-point pressure for carbon dioxide (CO_2) is 5.1 atm. At normal atmospheric pressure, solid CO_2 ("dry ice") undergoes sublimation; there is no liquid phase.

Line (b) in Fig. 18.24 represents compression at a constant temperature T_b. The material passes from vapor to liquid and then to solid at the points where line (b) crosses the vaporization curve and fusion curve, respectively. Line (d) shows constant-temperature compression at a lower temperature T_d; the material passes from vapor to solid at the point where line (d) crosses the sublimation curve.

We saw in the pV-diagram of Fig. 18.7 that a liquid-vapor phase transition occurs only when the temperature and pressure are less than those at the point at the top of the green shaded area labeled "Liquid-vapor phase equilibrium region." This point corresponds to the endpoint at the top of the vaporization curve in Fig. 18.24. It is called the **critical point,** and the corresponding values of p and T are called the critical pressure and temperature, p_c and T_c. A gas at a pressure *above* the critical pressure does not separate into two phases when it is cooled at constant pressure (along a horizontal line above the critical point in Fig. 18.24). Instead, its properties change gradually and continuously from those we ordinarily associate with a gas (low density, large compressibility) to those of a liquid (high density, small compressibility) *without a phase transition.*

TABLE 18.3 Triple-Point Data

Substance	Temperature (K)	Pressure (Pa)
Hydrogen	13.80	0.0704×10^5
Deuterium	18.63	0.171×10^5
Neon	24.56	0.432×10^5
Nitrogen	63.18	0.125×10^5
Oxygen	54.36	0.00152×10^5
Ammonia	195.40	0.0607×10^5
Carbon dioxide	216.55	5.17×10^5
Sulfur dioxide	197.68	0.00167×10^5
Water	273.16	0.00610×10^5

You can understand this by thinking about liquid-phase transitions at successively higher points on the vaporization curve. As we approach the critical point, the *differences* in physical properties (such as density and compressibility) between the liquid and vapor phases become smaller. Exactly *at* the critical point they all become zero, and at this point the distinction between liquid and vapor disappears. The heat of vaporization also grows smaller as we approach the critical point, and it too becomes zero at the critical point.

For nearly all familiar materials the critical pressures are much greater than atmospheric pressure, so we don't observe this behavior in everyday life. For example, the critical point for water is at 647.4 K and 221.2×10^5 Pa (about 218 atm or 3210 psi). But high-pressure steam boilers in electric generating plants regularly run at pressures and temperatures well above the critical point.

Many substances can exist in more than one solid phase. A familiar example is carbon, which exists as noncrystalline soot and crystalline graphite and diamond. Water is another example; more than a dozen types of ice, differing in crystal structure and physical properties, have been observed at very high pressures.

pVT-Surfaces

We remarked in Section 18.1 that for any material, it can be useful to represent the equation of state as a surface in a three-dimensional space with coordinates p, V, and T. **Figure 18.26** shows a typical pVT-surface. The light lines represent pV-isotherms; projecting them onto the pV-plane gives a diagram similar to Fig. 18.7. The pV-isotherms represent contour lines on the pVT-surface, just as contour lines on a topographic map represent the elevation (the third dimension) at each point. The projections of the edges of the surface onto the pT-plane give the pT phase diagram of Fig. 18.24.

Line *abcdef* in Fig. 18.26 represents constant-pressure heating, with melting along *bc* and vaporization along *de*. Note the volume changes that occur as T increases along this line. Line *ghjklm* corresponds to an isothermal (constant temperature) compression, with liquefaction along *hj* and solidification along *kl*. Between these, segments *gh* and *jk* represent isothermal compression with increase in pressure; the pressure increases are much greater in the liquid region *jk* and the solid region *lm* than in the vapor region *gh*. Finally, line *nopq* represents isothermal solidification directly from vapor, as in the formation of snowflakes or frost.

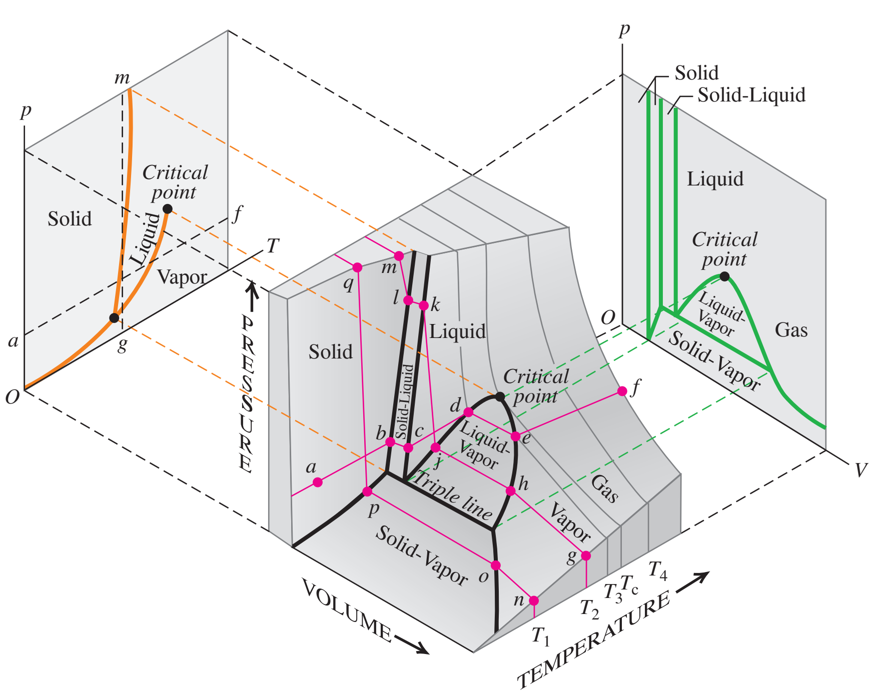


Figure **18.26** A pVT-surface for a substance that expands on melting. Projections of the boundaries on the surface onto the pT- and pV-planes are also shown.

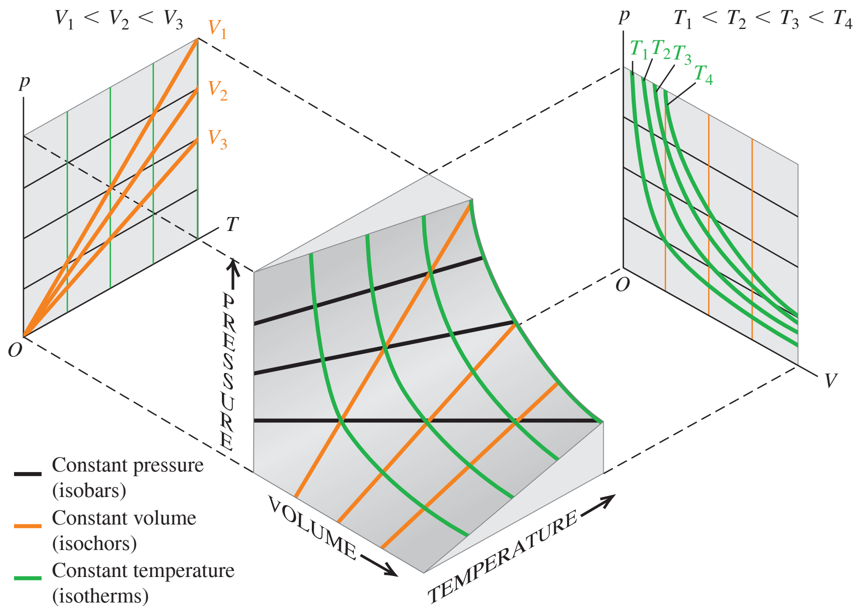


Figure **18.27** A *pVT*-surface for an ideal gas. At the left, each orange line corresponds to a certain constant volume; at the right, each green line corresponds to a certain constant temperature.

Figure 18.27 shows the much simpler *pVT*-surface for a substance that obeys the ideal-gas equation of state under all conditions. The projections of the constant-temperature curves onto the *pV*-plane correspond to the curves of Fig. 18.6, and the projections of the constant-volume curves onto the *pT*-plane show that pressure is directly proportional to absolute temperature. Figure 18.27 also shows the *isobars* (curves of constant pressure) and *isochors* (curves of constant volume) for an ideal gas.

TEST YOUR UNDERSTANDING OF SECTION 18.6 The average atmospheric pressure on Mars is 6.0×10^2 Pa. Could there be lakes of liquid water on the surface of Mars today? What about in the past, when the atmospheric pressure is thought to have been substantially greater?

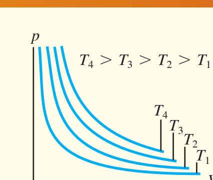
ANSWER

no, yes The triple-point pressure of water from Table 18.3 is 6.10×10^2 Pa. The present-day pressure on Mars is just less than this value, corresponding to the line labeled p_s in Fig. 18.24. Hence liquid water cannot exist on the present-day Martian surface, and there are no rivers or lakes. Planetary scientists conclude that liquid water could have existed and almost certainly did exist on Mars in the past, when the atmosphere was thicker.

CHAPTER 18 SUMMARY

Equations of state: The pressure p, volume V, and absolute temperature T of a given quantity of a substance are related by an equation of state. This relationship applies only for equilibrium states, in which p and T are uniform throughout the system. The ideal-gas equation of state, Eq. (18.3), involves the number of moles n and a constant R that is the same for all gases. (See Examples 18.1–18.4.)

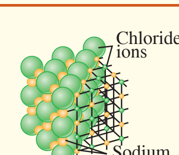
$$pV = nRT \quad (18.3)$$



Molecular properties of matter: The molar mass M of a pure substance is the mass per mole. The mass m_{total} of a quantity of substance equals M multiplied by the number of moles n. Avogadro's number N_A is the number of molecules in a mole. The mass m of an individual molecule is M divided by N_A. (See Example 18.5.)

$$m_{\text{total}} = nM \quad (18.2)$$

$$M = N_A m \quad (18.8)$$



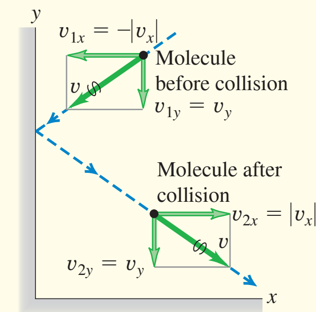
Kinetic-molecular model of an ideal gas: In an ideal gas, the total translational kinetic energy of the gas as a whole (K_{tr}) and the average translational kinetic energy per molecule $[\frac{1}{2}m(v^2)_{av}]$ are proportional to the absolute temperature T, and the root-mean-square speed of molecules is proportional to the square root of T. These expressions involve the Boltzmann constant $k = R/N_A$. (See Examples 18.6 and 18.7.) The mean free path λ of molecules in an ideal gas depends on the number of molecules per volume (N/V) and the molecular radius r. (See Example 18.8.)

$$K_{tr} = \tfrac{3}{2}nRT \qquad (18.14)$$

$$\tfrac{1}{2}m(v^2)_{av} = \tfrac{3}{2}kT \qquad (18.16)$$

$$v_{rms} = \sqrt{(v^2)_{av}} = \sqrt{\frac{3kT}{m}} \qquad (18.19)$$

$$\lambda = vt_{mean} = \frac{V}{4\pi\sqrt{2}r^2N} \qquad (18.21)$$

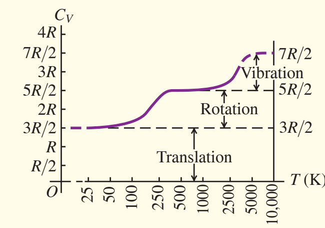


Heat capacities: The molar heat capacity at constant volume C_V is a simple multiple of the gas constant R for certain idealized cases: an ideal monatomic gas [Eq. (18.25)]; an ideal diatomic gas including rotational energy [Eq. (18.26)]; and an ideal monatomic solid [Eq. (18.28)]. Many real systems are approximated well by these idealizations.

$$C_V = \tfrac{3}{2}R \text{ (monatomic gas)} \qquad (18.25)$$

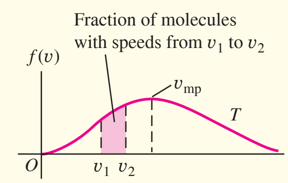
$$C_V = \tfrac{5}{2}R \text{ (diatomic gas)} \qquad (18.26)$$

$$C_V = 3R \text{ (monatomic solid)} \qquad (18.28)$$

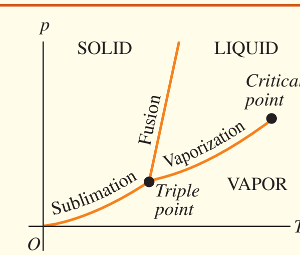


Molecular speeds: The speeds of molecules in an ideal gas are distributed according to the Maxwell–Boltzmann distribution $f(v)$. The quantity $f(v)\,dv$ describes what fraction of the molecules have speeds between v and $v + dv$.

$$f(v) = 4\pi\left(\frac{m}{2\pi kT}\right)^{3/2} v^2 e^{-mv^2/2kT} \qquad (18.32)$$



Phases of matter: Ordinary matter exists in the solid, liquid, and gas phases. A phase diagram shows conditions under which two phases can coexist in phase equilibrium. All three phases can coexist at the triple point. The vaporization curve ends at the critical point, above which the distinction between the liquid and gas phases disappears.



Chapter 18 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 18.1, 18.2, 18.3, and 18.4 (Section 18.1) before attempting these problems.

VP18.4.1 When the temperature is 30.0°C, the pressure of the air inside a bicycle tire of fixed volume 1.40×10^{-3} m^3 is 5.00×10^5 Pa. (a) What will be the pressure inside the tire when the temperature drops to 10.0°C? (b) How many moles of air are inside the tire?

VP18.4.2 When a research balloon is released at sea level, where the temperature is 15.0°C and the atmospheric pressure is 1.01×10^5 Pa, the helium in it has volume 13.0 m^3. (a) When the balloon reaches an altitude of 32.0 km, where the temperature is −44.5°C and the pressure is 868 Pa, what is the volume of the helium? (b) If this balloon is spherical, how many times larger is its radius at 32.0 km than at sea level?

VP18.4.3 The dwarf planet Pluto has a very thin atmosphere made up almost entirely of nitrogen (N_2, molar mass 2.8×10^{-2} kg/mol). At Pluto's surface the temperature is 42 K and the atmospheric pressure is 1.0 Pa. At the surface, (a) how many moles of gas are there per cubic meter of atmosphere, and (b) what is the density of the atmosphere in kg/m^3? (For comparison, the values are 42 mol/m^3 and 1.2 kg/m^3 at the earth's surface.)

VP18.4.4 When the pressure on n moles of helium gas is suddenly changed from an initial value of p_1 to a final value of p_2, the density of the gas changes from its initial value of ρ_1 to a final value of $\rho_2 = \rho_1(p_2/p_1)^{3/5}$. (a) If the initial absolute temperature of the gas is T_1, what is its final absolute temperature T_2 in terms of T_1, p_1, and p_2? (b) If the final pressure is 0.500 times the initial pressure, what are the ratio of the final density to the initial density and the ratio of the final temperature to the initial temperature? (c) Repeat part (b) if the final pressure is 2.00 times the initial pressure.

Be sure to review EXAMPLES 18.6 and 18.7 (Section 18.3) before attempting these problems.

VP18.7.1 At what temperature (in °C) is the rms speed of helium atoms (molar mass 4.00 g/mol) the same as the rms speed of nitrogen molecules (molar mass 28.0 g/mol) at 20.0°C? (Note that helium remains a gas at temperatures above −269°C.)

Continued

VP18.7.2 For the first 380,000 years after the Big Bang, the temperature of the matter in the universe was too high for nuclei and electrons to form atoms. The first hydrogen atoms (mass per atom 1.67×10^{-27} kg) did not form until the temperature had dropped to about 3000 K. (a) What was the average translational kinetic energy of hydrogen atoms when the temperature was 3.0×10^3 K? (b) What was the rms speed of these atoms?

VP18.7.3 The air in a room with dimensions 5.0 m × 5.0 m × 2.4 m is at temperature 20.0°C and pressure 1.00 atm = 1.01×10^5 Pa. (a) How many air molecules are in this room? (b) What is the total translational kinetic energy of these molecules? (c) How fast would a car of mass 1.5×10^3 kg have to move to have the same translational kinetic energy?

VP18.7.4 (a) What is the average value of the first 10 integers (1 through 10)? (b) What is the average value of the squares of the first 10 integers? (c) What is the rms value of the first 10 integers?

Be sure to review EXAMPLE 18.8 (Section 18.3) before attempting these problems.

VP18.8.1 At the surface of Mars the atmosphere has average pressure 6.0×10^2 Pa and average temperature −63°C. (a) What is the mean free path of atmospheric molecules (assumed to be spheres of radius 2.0×10^{-10} m) at the surface of Mars? (b) How many times greater is your answer in part (a) than the mean free path of atmospheric molecules at the earth's surface, where the average pressure is 1.01×10^5 Pa and the average temperature is 15°C?

VP18.8.2 A cubical box 1.00 m on a side contains air at 20.0°C. (a) What would the pressure inside the box have to be in order for the mean free path of air molecules (assumed to be spheres of radius 2.0×10^{-10} m) to be 1.00 m, so that a typical molecule suffers no collisions with other molecules as it travels the width of the box? (b) How many moles of air would be inside the box at the pressure calculated in part (a)? (At 1 atm pressure, this box would contain 41.4 mol of air.)

VP18.8.3 A cylinder for storing helium (molar mass 4.00 g/mol) has an interior volume of 50.0 L = 5.00×10^{-2} m^3. The cylinder holds 4.00×10^2 mol of helium under pressure at 27.0°C. If you assume the helium atoms are spheres of radius 3.1×10^{-11} m, what are (a) the mean free path of a helium atom in the tank and (b) the mean free time for a helium atom moving at the rms speed?

VP18.8.4 A gas with molecules of radius r and mass per molecule m is at temperature T and pressure p. (a) Write an expression for the mean free time for a molecule moving at the rms speed for this gas. (b) Which single change would have the greatest effect on the mean free time: doubling the radius r, doubling the pressure p, or doubling the temperature T?

BRIDGING PROBLEM Gas on Jupiter's Moon Europa

An astronaut visiting Jupiter's satellite Europa leaves a canister of 1.20 mol of nitrogen gas (28.0 g/mol) at 25.0°C on the satellite's surface. Europa has no significant atmosphere, and the acceleration due to gravity at its surface is 1.30 m/s^2. The canister springs a leak, allowing molecules to escape from a small hole. (a) What is the maximum height (in km) above Europa's surface that is reached by a nitrogen molecule whose speed equals the rms speed? Assume that the molecule is shot straight up out of the hole in the canister, and ignore the variation in g with altitude. (b) The escape speed from Europa is 2025 m/s. Can any of the nitrogen molecules escape from Europa and into space?

SOLUTION GUIDE

IDENTIFY and SET UP

1. Sketch the situation, showing all relevant dimensions.
2. List the unknown quantities, and decide which are the target variables.
3. How will you find the rms speed of the nitrogen molecules? What principle will you use to find the maximum height that a molecule with this speed can reach?
4. Does the rms molecular speed in the gas represent the maximum molecular speed? If not, what is the maximum speed?

EXECUTE

5. Solve for the rms speed. Use this to calculate the maximum height that a molecule with this speed can reach.
6. Use your result from step 5 to answer the question in part (b).

EVALUATE

7. Do your results depend on the amount of gas in the container? Why or why not?
8. How would your results from steps 5 and 6 be affected if the gas cylinder were instead left on Jupiter's satellite Ganymede, which has higher surface gravity than Europa and a higher escape speed? Like Europa, Ganymede has no significant atmosphere.

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q18.1 Section 18.1 states that ordinarily, pressure, volume, and temperature cannot change individually without one affecting the others. Yet when a liquid evaporates, its volume changes, even though its pressure and temperature are constant. Is this inconsistent? Why or why not?

Q18.2 In the ideal-gas equation, could an equivalent Celsius temperature be used instead of the Kelvin one if an appropriate numerical value of the constant R is used? Why or why not?

Q18.3 When a car is driven some distance, the air pressure in the tires increases. Why? Should you let out some air to reduce the pressure? Why or why not?

Q18.4 The coolant in an automobile radiator is kept at a pressure higher than atmospheric pressure. Why is this desirable? The radiator cap will release coolant when the gauge pressure of the coolant reaches a certain value, typically 15 lb/in.^{2} or so. Why not just seal the system completely?

Q18.5 Unwrapped food placed in a freezer experiences dehydration, known as "freezer burn." Why?

Q18.6 A group of students drove from their university (near sea level) up into the mountains for a skiing weekend. Upon arriving at the slopes, they discovered that the bags of potato chips they had brought for snacks had all burst open. What caused this to happen?

Q18.7 The derivation of the ideal-gas equation included the assumption that the number of molecules is very large, so that we could compute the average force due to many collisions. However, the ideal-gas equation holds accurately only at low pressures, where the molecules are few and far between. Is this inconsistent? Why or why not?

Q18.8 A rigid, perfectly insulated container has a membrane dividing its volume in half. One side contains a gas at an absolute temperature T_0 and pressure p_0, while the other half is completely empty. Suddenly a small hole develops in the membrane, allowing the gas to leak out into the other half until it eventually occupies twice its original volume. In terms of T_0 and p_0, what will be the new temperature and pressure of the gas when it is distributed equally in both halves of the container? Explain your reasoning.

Q18.9 (a) Which has more atoms: a kilogram of hydrogen or a kilogram of lead? Which has more mass? (b) Which has more atoms: a mole of hydrogen or a mole of lead? Which has more mass? Explain your reasoning.

Q18.10 Use the concepts of the kinetic-molecular model to explain: (a) why the pressure of a gas in a rigid container increases as heat is added to the gas and (b) why the pressure of a gas increases as we compress it, even if we do not change its temperature.

Q18.11 The proportions of various gases in the earth's atmosphere change somewhat with altitude. Would you expect the proportion of oxygen at high altitude to be greater or less than at sea level compared to the proportion of nitrogen? Why?

Q18.12 Comment on the following statement: *When two gases are mixed, if they are to be in thermal equilibrium, they must have the same average molecular speed.* Is the statement correct? Why or why not?

Q18.13 The kinetic-molecular model contains a hidden assumption about the temperature of the container walls. What is this assumption? What would happen if this assumption were not valid?

Q18.14 The temperature of an ideal gas is directly proportional to the average kinetic energy of its molecules. If a container of ideal gas is moving past you at 2000 m/s, is the temperature of the gas higher than if the container was at rest? Explain your reasoning.

Q18.15 If the pressure of an ideal monatomic gas is increased while the number of moles is kept constant, what happens to the average translational kinetic energy of one atom of the gas? Is it possible to change *both* the volume and the pressure of an ideal gas and keep the average translational kinetic energy of the atoms constant? Explain.

Q18.16 In deriving the ideal-gas equation from the kinetic-molecular model, we ignored potential energy due to the earth's gravity. Is this omission justified? Why or why not?

Q18.17 Imagine a special air filter placed in a window of a house. The tiny holes in the filter allow only air molecules moving faster than a certain speed to exit the house, and allow only air molecules moving slower than that speed to enter the house from outside. What effect would this filter have on the temperature inside the house? (It turns out that the second law of thermodynamics—which we'll discuss in Chapter 20—tells us that such a wonderful air filter would be impossible to make.)

Q18.18 A gas storage tank has a small leak. The pressure in the tank drops more quickly if the gas is hydrogen or helium than if it is oxygen. Why?

Q18.19 Consider two specimens of ideal gas at the same temperature. Specimen A has the same total mass as specimen B, but the molecules in specimen A have greater molar mass than they do in specimen B. In which specimen is the total kinetic energy of the gas greater? Does your answer depend on the molecular structure of the gases? Why or why not?

Q18.20 The temperature of an ideal monatomic gas is increased from 25°C to 50°C. Does the average translational kinetic energy of each gas atom double? Explain. If your answer is no, what would the final temperature be if the average translational kinetic energy was doubled?

Q18.21 If the root-mean-square speed of the atoms of an ideal gas is to be doubled, by what factor must the Kelvin temperature of the gas be increased? Explain.

Q18.22 (a) If you apply the same amount of heat to 1.00 mol of an ideal monatomic gas and 1.00 mol of an ideal diatomic gas, which one (if any) will increase more in temperature? (b) Physically, *why* do diatomic gases have a greater molar heat capacity than monatomic gases?

Q18.23 The discussion in Section 18.4 concluded that all ideal monatomic gases have the same heat capacity C_V. Does this mean that it takes the same amount of heat to raise the temperature of 1.0 g of each one by 1.0 K? Explain your reasoning.

Q18.24 In a gas that contains N molecules, is it accurate to say that the number of molecules with speed v is equal to $f(v)$? Is it accurate to say that this number is given by $Nf(v)$? Explain your answers.

Q18.25 The atmosphere of the planet Mars is 95.3% carbon dioxide (CO_2) and about 0.03% water vapor. The atmospheric pressure is only about 600 Pa, and the surface temperature varies from -30°C to -100°C. The polar ice caps contain both CO_2 ice and water ice. Could there be *liquid* CO_2 on the surface of Mars? Could there be liquid water? Why or why not?

Q18.26 A beaker of water at room temperature is placed in an enclosure, and the air pressure in the enclosure is slowly reduced. When the air pressure is reduced sufficiently, the water begins to boil. The temperature of the water does not rise when it boils; in fact, the temperature *drops* slightly. Explain these phenomena.

Q18.27 When you stand on ice with skates, the blades on the bottom have a low surface area, and so they pressurize the ice beneath the blades. When ice is pressurized, is the melting temperature greater than or less than 0°C? Explain.

Q18.28 Hydrothermal vents are openings in the ocean floor that discharge very hot water. The water emerging from one such vent off the Oregon coast, 2400 m below the surface, is at 279°C. Despite its high temperature, the water doesn't boil. Why not?

Q18.29 The dark areas on the moon's surface are called *maria*, Latin for "seas," and were once thought to be bodies of water. In fact, the maria are not "seas" at all, but plains of solidified lava. Given that there is no atmosphere on the moon, how can you explain the absence of liquid water on the moon's surface?

Q18.30 In addition to the normal cooking directions printed on the back of a box of rice, there are also "high-altitude directions." The only difference is that the "high-altitude directions" suggest increasing the cooking time and using a greater volume of boiling water in which to cook the rice. Why should the directions depend on the altitude in this way?

EXERCISES

Section 18.1 Equations of State

18.1 • A 20.0 L tank contains 4.86×10^{-4} kg of helium at 18.0°C. The molar mass of helium is 4.00 g/mol. (a) How many moles of helium are in the tank? (b) What is the pressure in the tank, in pascals and in atmospheres?

18.2 ••• Helium gas with a volume of 3.20 L, under a pressure of 0.180 atm and at 41.0°C, is warmed until both pressure and volume are doubled. (a) What is the final temperature? (b) How many grams of helium are there? The molar mass of helium is 4.00 g/mol.

18.3 • A cylindrical tank has a tight-fitting piston that allows the volume of the tank to be changed. The tank originally contains 0.110 m^3 of air at a pressure of 0.355 atm. The piston is slowly pulled out until the volume of the gas is increased to 0.390 m^3. If the temperature remains constant, what is the final value of the pressure?

18.4 • A 3.00 L tank contains air at 3.00 atm and 20.0°C. The tank is sealed and cooled until the pressure is 1.00 atm. (a) What is the temperature then in degrees Celsius? Assume that the volume of the tank is constant. (b) If the temperature is kept at the value found in part (a) and the gas is compressed, what is the volume when the pressure again becomes 3.00 atm?

18.5 • The discussion following Eq. (18.7) gives the constants in the van der Waals equation for CO_2 gas. It also says that at STP the van der Waals equation gives only a small (0.5%) correction in the ideal-gas equation. Consider 1 mole of CO_2 gas at $T = 273.0$ K and a volume of 4.48×10^{-4} m^3. (a) What is the pressure of the gas calculated by the ideal-gas equation? (b) What does the van der Waals equation give for the pressure? What is the percentage difference from the ideal-gas result?

18.6 •• You have several identical balloons. You experimentally determine that a balloon will break if its volume exceeds 0.900 L.The pressure of the gas inside the balloon equals air pressure (1.00 atm). (a) If the air inside the balloon is at a constant 22.0°C and behaves as an ideal gas, what mass of air can you blow into one of the balloons before it bursts? (b) Repeat part (a) if the gas is helium rather than air.

18.7 •• A Jaguar XK8 convertible has an eight-cylinder engine. At the beginning of its compression stroke, one of the cylinders contains 499 cm^3 of air at atmospheric pressure (1.01×10^5 Pa) and a temperature of 27.0°C. At the end of the stroke, the air has been compressed to a volume of 46.2 cm^3 and the gauge pressure has increased to 2.72×10^6 Pa. Compute the final temperature.

18.8 •• A welder using a tank of volume 0.0750 m^3 fills it with oxygen (molar mass 32.0 g/mol) at a gauge pressure of 3.00×10^5 Pa and temperature of 37.0°C. The tank has a small leak, and in time some of the oxygen leaks out. On a day when the temperature is 22.0°C, the gauge pressure of the oxygen in the tank is 1.80×10^5 Pa. Find (a) the initial mass of oxygen and (b) the mass of oxygen that has leaked out.

18.9 •• A large cylindrical tank contains 0.750 m^3 of nitrogen gas at 27°C and 7.50×10^3 Pa (absolute pressure). The tank has a tight-fitting piston that allows the volume to be changed. What will be the pressure if the volume is decreased to 0.410 m^3 and the temperature is increased to 157°C?

18.10 • An empty cylindrical canister 1.50 m long and 90.0 cm in diameter is to be filled with pure oxygen at 22.0°C to store in a space station. To hold as much gas as possible, the absolute pressure of the oxygen will be 21.0 atm. The molar mass of oxygen is 32.0 g/mol. (a) How many moles of oxygen does this canister hold? (b) For someone lifting this canister, by how many kilograms does this gas increase the mass to be lifted?

18.11 • The gas inside a balloon will always have a pressure nearly equal to atmospheric pressure, since that is the pressure applied to the outside of the balloon. You fill a balloon with helium (a nearly ideal gas) to a volume of 0.600 L at 19.0°C. What is the volume of the balloon if you cool it to the boiling point of liquid nitrogen (77.3 K)?

18.12 • An ideal gas has a density of 1.33×10^{-6} g/cm^3 at 1.00×10^{-3} atm and 20.0°C. Identify the gas.

18.13 •• If a certain amount of ideal gas occupies a volume V at STP on earth, what would be its volume (in terms of V) on Venus, where the temperature is 1003°C and the pressure is 92 atm?

18.14 • A diver observes a bubble of air rising from the bottom of a lake (where the absolute pressure is 3.50 atm) to the surface (where the pressure is 1.00 atm). The temperature at the bottom is 4.0°C, and the temperature at the surface is 23.0°C. (a) What is the ratio of the volume of the bubble as it reaches the surface to its volume at the bottom? (b) Would it be safe for the diver to hold his breath while ascending from the bottom of the lake to the surface? Why or why not?

18.15 • How many moles of an ideal gas exert a gauge pressure of 0.876 atm in a volume of 5.43 L at a temperature of 22.2°C?

18.16 • Three moles of an ideal gas are in a rigid cubical box with sides of length 0.300 m. (a) What is the force that the gas exerts on each of the six sides of the box when the gas temperature is 20.0°C? (b) What is the force when the temperature of the gas is increased to 100.0°C?

18.17 •• (a) Calculate the mass of nitrogen present in a volume of 3000 cm^3 if the gas is at 22.0°C and the absolute pressure of 2.00×10^{-13} atm is a partial vacuum easily obtained in laboratories. (b) What is the density (in kg/m^3) of the N_2?

18.18 • At an altitude of 11,000 m (a typical cruising altitude for a jet airliner), the air temperature is -56.5°C and the air density is 0.364 kg/m^3. What is the pressure of the atmosphere at that altitude? (*Note:* The temperature at this altitude is not the same as at the surface of the earth, so the calculation of Example 18.4 in Section 18.1 doesn't apply.)

Section 18.2 Molecular Properties of Matter

18.19 • How many moles are in a 1.00 kg bottle of water? How many molecules? The molar mass of water is 18.0 g/mol.

18.20 • A large organic molecule has a mass of 1.41×10^{-21} kg. What is the molar mass of this compound?

18.21 •• Modern vacuum pumps make it easy to attain pressures of the order of 10^{-13} atm in the laboratory. Consider a volume of air and treat the air as an ideal gas. (a) At a pressure of 9.00×10^{-14} atm and an ordinary temperature of 300.0 K, how many molecules are present in a volume of 1.00 cm^3? (b) How many molecules would be present at the same temperature but at 1.00 atm instead?

18.22 •• The Lagoon Nebula (**Fig. E18.22**) is a cloud of hydrogen gas located 3900 light-years from the earth. The cloud is about 45 light-years in diameter and glows because of its high temperature of 7500 K. (The gas is raised to this temperature by the stars that lie within the nebula.) The cloud is also very thin; there are only 80 molecules per cubic centimeter. (a) Find the gas pressure (in atmospheres) in the Lagoon Nebula. Compare it to the laboratory pressure referred to in Exercise 18.21. (b) Science-fiction films sometimes show starships being buffeted by turbulence as they fly through gas clouds such as the Lagoon Nebula. Does this seem realistic? Why or why not?

Figure **E18.22**

18.23 •• **How Close Together Are Gas Molecules?** Consider an ideal gas at 27°C and 1.00 atm. To get some idea how close these molecules are to each other, on the average, imagine them to be uniformly spaced, with each molecule at the center of a small cube. (a) What is the length of an edge of each cube if adjacent cubes touch but do not overlap? (b) How does this distance compare with the diameter of a typical molecule? (c) How does their separation compare with the spacing of atoms in solids, which typically are about 0.3 nm apart?

Section 18.3 Kinetic-Molecular Model of an Ideal Gas

18.24 •• A container with rigid walls holds n moles of a monatomic ideal gas. In terms of n, how many moles of the gas must be removed from the container to double the pressure while also doubling the rms speed of the gas atoms?

18.25 • (a) What is the total translational kinetic energy of the air in an empty room that has dimensions 8.00 m × 12.00 m × 4.00 m if the air is treated as an ideal gas at 1.00 atm? (b) What is the speed of a 2000 kg automobile if its kinetic energy equals the translational kinetic energy calculated in part (a)?

18.26 • A flask contains a mixture of neon (Ne), krypton (Kr), and radon (Rn) gases. Compare (a) the average kinetic energies of the three types of atoms and (b) the root-mean-square speeds. (*Hint:* Appendix D shows the molar mass (in g/mol) of each element under the chemical symbol for that element.)

18.27 • A container with volume 80.0 cm^3 and rigid walls holds a monatomic ideal gas. To determine the number of gas atoms in the container, you measure the pressure p_{atm} of the gas in atmospheres as a function of the Celsius temperature T_C of the gas. You plot p_{atm} versus T_C and find that your data lie close to a straight line that has slope 1.10 atm/C°. What is your experimental result for the number of gas atoms?

18.28 • A container with volume 1.64 L is initially evacuated. Then it is filled with 0.226 g of N_2. Assume that the pressure of the gas is low enough for the gas to obey the ideal-gas law to a high degree of accuracy. If the root-mean-square speed of the gas molecules is 182 m/s, what is the pressure of the gas?

18.29 •• (a) A deuteron, ^2_1H, is the nucleus of a hydrogen isotope and consists of one proton and one neutron. The plasma of deuterons in a nuclear fusion reactor must be heated to about 300 million K. What is the rms speed of the deuterons? Is this a significant fraction of the speed of light in vacuum ($c = 3.0 \times 10^8$ m/s)? (b) What would the temperature of the plasma be if the deuterons had an rms speed equal to $0.10c$?

18.30 • **Martian Climate.** The atmosphere of Mars is mostly CO_2 (molar mass 44.0 g/mol) under a pressure of 650 Pa, which we shall assume remains constant. In many places the temperature varies from 0.0°C in summer to −100°C in winter. Over the course of a Martian year, what are the ranges of (a) the rms speeds of the CO_2 molecules and (b) the density (in mol/m^3) of the atmosphere?

18.31 •• Oxygen (O_2) has a molar mass of 32.0 g/mol. What is (a) the average translational kinetic energy of an oxygen molecule at a temperature of 300 K; (b) the average value of the square of its speed; (c) the root-mean-square speed; (d) the momentum of an oxygen molecule traveling at this speed? (e) Suppose an oxygen molecule traveling at this speed bounces back and forth between opposite sides of a cubical vessel 0.10 m on a side. What is the average force the molecule exerts on one of the walls of the container? (Assume that the molecule's velocity is perpendicular to the two sides that it strikes.) (f) What is the average force per unit area? (g) How many oxygen molecules traveling at this speed are necessary to produce an average pressure of 1 atm? (h) Compute the number of oxygen molecules that are contained in a vessel of this size at 300 K and atmospheric pressure. (i) Your answer for part (h) should be three times as large as the answer for part (g). Where does this discrepancy arise?

18.32 •• Calculate the mean free path of air molecules at 3.50×10^{-13} atm and 300 K. (This pressure is readily attainable in the laboratory; see Exercise 18.21.) As in Example 18.8, model the air molecules as spheres of radius 2.0×10^{-10} m.

18.33 •• At what temperature is the root-mean-square speed of nitrogen molecules equal to the root-mean-square speed of hydrogen molecules at 20.0°C? (*Hint:* Appendix D shows the molar mass (in g/mol) of each element under the chemical symbol for that element. The molar mass of H_2 is twice the molar mass of hydrogen atoms, and similarly for N_2.)

18.34 • Smoke particles in the air typically have masses on the order of 10^{-16} kg. The Brownian motion (rapid, irregular movement) of these particles, resulting from collisions with air molecules, can be observed with a microscope. (a) Find the root-mean-square speed of Brownian motion for a particle with a mass of 3.00×10^{-16} kg in air at 300 K. (b) Would the root-mean-square speed be different if the particle were in hydrogen gas at the same temperature? Explain.

Section 18.4 Heat Capacities

18.35 •• Three moles of helium gas (molar mass $M = 4.00$ g/mol) are in a rigid container that keeps the volume of the gas constant. Initially the rms speed of the gas atoms is 900 m/s. What is the rms speed of the gas atoms after 2400 J of heat energy is added to the gas?

18.36 •• A rigid container holds 4.00 mol of a monatomic ideal gas that has temperature 300 K. The initial pressure of the gas is 6.00×10^4 Pa. What is the pressure after 6000 J of heat energy is added to the gas?

18.37 • How much heat does it take to increase the temperature of 1.80 mol of an ideal gas by 50.0 K near room temperature if the gas is held at constant volume and is (a) diatomic; (b) monatomic?

18.38 •• Perfectly rigid containers each hold n moles of ideal gas, one being hydrogen (H_2) and the other being neon (Ne). If it takes 300 J of heat to increase the temperature of the hydrogen by 2.50°C, by how many degrees will the same amount of heat raise the temperature of the neon?

18.39 •• (a) Compute the specific heat at constant volume of nitrogen (N_2) gas, and compare it with the specific heat of liquid water. The molar mass of N_2 is 28.0 g/mol. (b) You warm 1.00 kg of water at a constant volume of 1.00 L from 20.0°C to 30.0°C in a kettle. For the same amount of heat, how many kilograms of 20.0°C air would you be able to warm to 30.0°C? What volume (in liters) would this air occupy at 20.0°C and a pressure of 1.00 atm? Make the simplifying assumption that air is 100% N_2.

18.40 •• (a) Calculate the specific heat at constant volume of water vapor, assuming the nonlinear triatomic molecule has three translational and three rotational degrees of freedom and that vibrational motion does not contribute. The molar mass of water is 18.0 g/mol. (b) The actual specific heat of water vapor at low pressures is about 2000 J/kg · K. Compare this with your calculation and comment on the actual role of vibrational motion.

Section 18.5 Molecular Speeds

18.41 • For diatomic carbon dioxide gas (CO_2, molar mass 44.0 g/mol) at $T = 300$ K, calculate (a) the most probable speed v_{mp}; (b) the average speed v_{av}; (c) the root-mean-square speed v_{rms}.

18.42 • For a gas of nitrogen molecules (N_2), what must the temperature be if 94.7% of all the molecules have speeds less than (a) 1500 m/s; (b) 1000 m/s; (c) 500 m/s? Use Table 18.2. The molar mass of N_2 is 28.0 g/mol.

18.43 • The speed of sound for an ideal gas is given by $v = \sqrt{\gamma RT/M}$ [Eq. (16.10)]. We'll see in Chapter 19 that for a monatomic ideal gas, $\gamma = 1.67$. (a) What is the ratio v_{rms}/v? (b) The average speed v_{av} of the gas atoms is given by Eq. (18.35). What is the ratio v_{av}/v?

18.44 • In 0.0345 mol of a monatomic ideal gas, how many of the atoms have speeds that are within 20% of the rms speed? (Use Table 18.2.)

Section 18.6 Phases of Matter

18.45 • Solid water (ice) is slowly warmed from a very low temperature. (a) What minimum external pressure p_1 must be applied to the solid if a melting phase transition is to be observed? Describe the sequence of phase transitions that occur if the applied pressure p is such that $p < p_1$. (b) Above a certain maximum pressure p_2, no boiling transition is observed. What is this pressure? Describe the sequence of phase transitions that occur if $p_1 < p < p_2$.

18.46 • **Meteorology.** The *vapor pressure* is the pressure of the vapor phase of a substance when it is in equilibrium with the solid or liquid phase of the substance. The *relative humidity* is the partial pressure of water vapor in the air divided by the vapor pressure of water at that same temperature, expressed as a percentage. The air is saturated when the humidity is 100%. (a) The vapor pressure of water at 20.0°C is 2.34×10^3 Pa. If the air temperature is 20.0°C and the relative humidity is 60%, what is the partial pressure of water vapor in the atmosphere (that is, the pressure due to water vapor alone)? (b) Under the conditions of part (a), what is the mass of water in 1.00 m^3 of air? (The molar mass of water is 18.0 g/mol. Assume that water vapor can be treated as an ideal gas.)

PROBLEMS

18.47 •• **CP** Imagine the sound made when a latex balloon with a diameter of 30 cm pops 2 m from your ear. Estimate the sound intensity level in decibels, using Table 16.2 as a guide. (a) Assuming the duration of the popping event was 100 ms, use the intensity of the popping sound to determine the average power of the pop and the energy released in the pop. This provides an estimate of the energy stored in the balloon prior to the pop. (b) Let T be the ratio of the energy stored in the stretched balloon to its surface area. Use your estimate of the stored energy and the size of the balloon to estimate the value of T. (c) The gauge pressure of a latex balloon depends on T and radius R according to $p_{gauge} = T/R$. Use this information to estimate the gauge pressure in the balloon.

18.48 • A physics lecture room at 1.00 atm and 27.0°C has a volume of 216 m^3. (a) Use the ideal-gas law to estimate the number of air molecules in the room. Assume that all of the air is N_2. Calculate (b) the particle density—that is, the number of N_2 molecules per cubic centimeter—and (c) the mass of the air in the room.

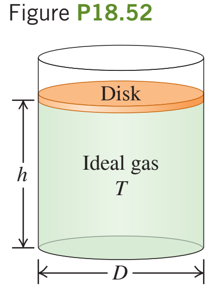
18.49 •• CP BIO **The Effect of Altitude on the Lungs.** (a) Calculate the *change* in air pressure you'll experience if you climb a 1000 m mountain, assuming for simplicity that the temperature and air density do not change over this distance and that they were 22°C and 1.2 kg/m^3, respectively, at the bottom of the mountain. (*Note:* The result of Example 18.4 doesn't apply, since the expression derived in that example accounts for the variation of air density with altitude and we are told to ignore that here.) (b) If you took a 0.50 L breath at the foot of the mountain and managed to hold it until you reached the top, what would be the volume of this breath when you exhaled it there?

18.50 •• CP BIO **The Bends.** If deep-sea divers rise to the surface too quickly, nitrogen bubbles in their blood can expand and prove fatal. This phenomenon is known as the *bends.* If a scuba diver rises quickly from a depth of 25 m in Lake Michigan (which is fresh water), what will be the volume at the surface of an N_2 bubble that occupied 1.0 mm^3 in his blood at the lower depth? Does it seem that this difference is large enough to be a problem? (Assume that the pressure difference is due to only the changing water pressure, not to any temperature difference. This assumption is reasonable, since we are warm-blooded creatures.)

18.51 ••• CP A hot-air balloon stays aloft because hot air at atmospheric pressure is less dense than cooler air at the same pressure. If the volume of the balloon is 500.0 m^3 and the surrounding air is at 15.0°C, what must the temperature of the air in the balloon be for it to lift a total load of 290 kg (in addition to the mass of the hot air)? The density of air at 15.0°C and atmospheric pressure is 1.23 kg/m^3.

18.52 •• In an evacuated enclosure, a vertical cylindrical tank of diameter D is sealed by a 3.00 kg circular disk that can move up and down without friction. Beneath the disk is a quantity of ideal gas at temperature T in the cylinder (**Fig. P18.52**). Initially the disk is at rest at a distance of $h = 4.00$ m above the bottom of the tank. When a lead brick of mass 9.00 kg is gently placed on the disk, the disk moves downward. If the temperature of the gas is kept constant and no gas escapes from the tank, what distance above the bottom of the tank is the disk when it again comes to rest?

Figure P18.52



18.53 ••• A cylinder 1.00 m tall with inside diameter 0.120 m is used to hold propane gas (molar mass 44.1 g/mol) for use in a barbecue. It is initially filled with gas until the gauge pressure is 1.30×10^6 Pa at 22.0°C. The temperature of the gas remains constant as it is partially emptied out of the tank, until the gauge pressure is 3.40×10^5 Pa. Calculate the mass of propane that has been used.

18.54 • CP During a test dive in 1939, prior to being accepted by the U.S. Navy, the submarine *Squalus* sank at a point where the depth of water was 73.0 m. The temperature was 27.0°C at the surface and 7.0°C at the bottom. The density of seawater is 1030 kg/m^3. (a) A diving bell was used to rescue 33 trapped crewmen from the *Squalus.* The diving bell was in the form of a circular cylinder 2.30 m high, open at the bottom and closed at the top. When the diving bell was lowered to the bottom of the sea, to what height did water rise within the diving bell? (*Hint:* Ignore the relatively small variation in water pressure between the bottom of the bell and the surface of the water within the bell.) (b) At what gauge pressure must compressed air have been supplied to the bell while on the bottom to expel all the water from it?

18.55 •• A parcel of air over a campfire feels an upward buoyant force because the heated air is less dense than the surrounding air. By estimating the acceleration of the air immediately above a fire, one can estimate the fire's temperature. The mass of a volume V of air is nM_{air}, where n is the number of moles of air molecules in the volume and M_{air} is the molar mass of air. The net upward force on a parcel of air above a fire is roughly given by $(m_{out} - m_{in})g$, where m_{out} is the mass of a volume of ambient air and m_{in} is the mass of a similar volume of air in the hot zone. (a) Use the ideal-gas law, along with the knowledge that the pressure of the air above the fire is the same as that of the ambient air, to derive an expression for the acceleration a of an air parcel as a function of (T_{out}/T_{in}), where T_{in} is the absolute temperature of the air above the fire and T_{out} is the absolute temperature of the ambient air. (b) Rearrange your formula from part (a) to obtain an expression for T_{in} as a function of T_{out} and a. (c) Based on your experience with campfires, estimate the acceleration of the air above the fire by comparing in your mind the upward trajectory of sparks with the acceleration of falling objects. Thus you can estimate a as a multiple of g. (d) Assuming an ambient temperature of 15°C, use your formula and your estimate of a to estimate the temperature of the fire.

18.56 • **Pressure on Venus.** At the surface of Venus the average temperature is a balmy 460°C due to the greenhouse effect (global warming!), the pressure is 92 earth-atmospheres, and the acceleration due to gravity is $0.894g_{earth}$. The atmosphere is nearly all CO_2 (molar mass 44.0 g/mol), and the temperature remains remarkably constant. Assume that the temperature does not change with altitude. (a) What is the atmospheric pressure 1.00 km above the surface of Venus? Express your answer in Venus-atmospheres and earth-atmospheres. (b) What is the root-mean-square speed of the CO_2 molecules at the surface of Venus and at an altitude of 1.00 km?

18.57 ••• CP CALC A cylindrical diving bell has a radius of 750 cm and a height of 2.50 m. The bell includes a top compartment that holds an undersea adventurer. A bottom compartment separated from the top by a sturdy grating holds a tank of compressed air with a valve to release air into the bell, a second valve that can release air from the bell into the sea, a third valve that regulates the entry of seawater for ballast, a pump that removes the ballast to increase buoyancy, and an electric heater that maintains a constant temperature of 20.0°C. The total mass of the bell and all of its apparatuses is 4350 kg. The density of seawater is 1025 kg/m^3. (a) An 80.0 kg adventurer enters the bell. How many liters of seawater should be moved into the bell so that it is neutrally buoyant? (b) By carefully regulating ballast, the bell is made to descend into the sea at a rate of 1.0 m/s. Compressed air is released from the tank to raise the pressure in the bell to match the pressure of the seawater outside the bell. As the bell descends, at what rate should air be released through the first valve? (*Hint:* Derive an expression for the number of moles of air in the bell n as a function of depth y; then differentiate this to obtain dn/dt as a function of dy/dt.) (c) If the compressed air tank is a fully loaded, specially designed, 600 ft^3 tank, which means it contains that volume of air at standard temperature and pressure (0°C and 1 atm), how deep can the bell descend?

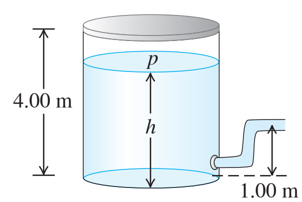
18.58 •• A flask with a volume of 1.50 L, provided with a stopcock, contains ethane gas (C_2H_6) at 300 K and atmospheric pressure (1.013×10^5 Pa). The molar mass of ethane is 30.1 g/mol. The system is warmed to a temperature of 550 K, with the stopcock open to the atmosphere. The stopcock is then closed, and the flask is cooled to its original temperature. (a) What is the final pressure of the ethane in the flask? (b) How many grams of ethane remain in the flask?

18.59 •• CP A balloon of volume 750 m^3 is to be filled with hydrogen at atmospheric pressure (1.01×10^5 Pa). (a) If the hydrogen is stored in cylinders with volumes of 1.90 m^3 at a gauge pressure of 1.20×10^6 Pa, how many cylinders are required? Assume that the temperature of the hydrogen remains constant. (b) What is the total weight (in addition to the weight of the gas) that can be supported by the balloon if both the gas in the balloon and the surrounding air are at 15.0°C? The molar mass of hydrogen (H_2) is 2.02 g/mol. The density of air at 15.0°C and atmospheric pressure is 1.23 kg/m^3. See Chapter 12 for a discussion of buoyancy. (c) What weight could be supported if the balloon were filled with helium (molar mass 4.00 g/mol) instead of hydrogen, again at 15.0°C?

18.60 •• A vertical cylindrical tank contains 1.80 mol of an ideal gas under a pressure of 0.300 atm at 20.0°C. The round part of the tank has a radius of 10.0 cm, and the gas is supporting a piston that can move up and down in the cylinder without friction. There is a vacuum above the piston. (a) What is the mass of this piston? (b) How tall is the column of gas that is supporting the piston?

18.61 •• CP A large tank of water has a hose connected to it (**Fig. P18.61**). The tank is sealed at the top and has compressed air between the water surface and the top. When the water height h has the value 3.50 m, the absolute pressure p of the compressed air is 4.20×10^5 Pa. Assume that the air above the water expands at constant temperature, and take the atmospheric pressure to be 1.00×10^5 Pa. (a) What is the speed with which water flows out of the hose when $h = 3.50$ m? (b) As water flows out of the tank, h decreases. Calculate the speed of flow for $h = 3.00$ m and for $h = 2.00$ m. (c) At what value of h does the flow stop?

Figure **P18.61**



18.62 •• CP A light, plastic sphere with mass $m = 9.00$ g and density $\rho = 4.00$ kg/m^3 is suspended in air by thread of negligible mass. (a) What is the tension T in the thread if the air is at 5.00°C and $p = 1.00$ atm? The molar mass of air is 28.8 g/mol. (b) How much does the tension in the thread change if the temperature of the gas is increased to 35.0°C? Ignore the change in volume of the plastic sphere when the temperature is changed.

18.63 •• BIO **How Many Atoms Are You?** Estimate the number of atoms in the body of a 50 kg physics student. Note that the human body is mostly water, which has molar mass 18.0 g/mol, and that each water molecule contains three atoms.

18.64 •• BIO A person at rest inhales 0.50 L of air with each breath at a pressure of 1.00 atm and a temperature of 20.0°C. The inhaled air is 21.0% oxygen. (a) How many oxygen molecules does this person inhale with each breath? (b) Suppose this person is now resting at an elevation of 2000 m but the temperature is still 20.0°C. Assuming that the oxygen percentage and volume per inhalation are the same as stated above, how many oxygen molecules does this person now inhale with each breath? (c) Given that the body still requires the same number of oxygen molecules per second as at sea level to maintain its functions, explain why some people report "shortness of breath" at high elevations.

18.65 •• You blow up a spherical balloon to a diameter of 50.0 cm until the absolute pressure inside is 1.25 atm and the temperature is 22.0°C. Assume that all the gas is N_2, of molar mass 28.0 g/mol. (a) Find the mass of a single N_2 molecule. (b) How much translational kinetic energy does an average N_2 molecule have? (c) How many N_2 molecules are in this balloon? (d) What is the *total* translational kinetic energy of all the molecules in the balloon?

18.66 • The size of an oxygen molecule is about 2.0×10^{-10} m. Make a rough estimate of the pressure at which the finite volume of the molecules should cause noticeable deviations from ideal-gas behavior at ordinary temperatures ($T = 300$ K).

18.67 •• CP CALC **The Lennard-Jones Potential.** A commonly used potential-energy function for the interaction of two molecules (see Fig. 18.8) is the Lennard-Jones 6-12 potential:

$$U(r) = U_0\left[\left(\frac{R_0}{r}\right)^{12} - 2\left(\frac{R_0}{r}\right)^{6}\right]$$

where r is the distance between the centers of the molecules and U_0 and R_0 are positive constants. The corresponding force $F(r)$ is given in Eq. (14.26). (a) Graph $U(r)$ and $F(r)$ versus r. (b) Let r_1 be the value of r at which $U(r) = 0$, and let r_2 be the value of r at which $F(r) = 0$. Show the locations of r_1 and r_2 on your graphs of $U(r)$ and $F(r)$. Which of these values represents the equilibrium separation between the molecules? (c) Find the values of r_1 and r_2 in terms of R_0, and find the ratio r_1/r_2. (d) If the molecules are located a distance r_2 apart [as calculated in part (c)], how much work must be done to pull them apart so that $r \rightarrow \infty$?

18.68 • (a) What is the total random translational kinetic energy of 5.00 L of hydrogen gas (molar mass 2.016 g/mol) with pressure 1.01×10^5 Pa and temperature 300 K? (*Hint:* Use the procedure of Problem 18.65 as a guide.) (b) If the tank containing the gas is placed on a swift jet moving at 300.0 m/s, by what percentage is the *total* kinetic energy of the gas increased? (c) Since the kinetic energy of the gas molecules is greater when it is on the jet, does this mean that its temperature has gone up? Explain.

18.69 •• It is possible to make crystalline solids that are only one layer of atoms thick. Such "two-dimensional" crystals can be created by depositing atoms on a very flat surface. (a) If the atoms in such a two-dimensional crystal can move only within the plane of the crystal, what will be its molar heat capacity near room temperature? Give your answer as a multiple of R and in J/mol·K. (b) At very low temperatures, will the molar heat capacity of a two-dimensional crystal be greater than, less than, or equal to the result you found in part (a)? Explain why.

18.70 • **Hydrogen on the Sun.** The surface of the sun has a temperature of about 5800 K and consists largely of hydrogen atoms. (a) Find the rms speed of a hydrogen atom at this temperature. (The mass of a single hydrogen atom is 1.67×10^{-27} kg.) (b) The escape speed for a particle to leave the gravitational influence of the sun is given by $(2GM/R)^{1/2}$, where M is the sun's mass, R its radius, and G the gravitational constant (see Example 13.5 of Section 13.3). Use Appendix F to calculate this escape speed. (c) Can appreciable quantities of hydrogen escape from the sun? Can *any* hydrogen escape? Explain.

18.71 •• CP (a) Show that a projectile with mass m can "escape" from the surface of a planet if it is launched vertically upward with a kinetic energy greater than mgR_p, where g is the acceleration due to gravity at the planet's surface and R_p is the planet's radius. Ignore air resistance. (See Problem 18.70.) (b) If the planet in question is the earth, at what temperature does the average translational kinetic energy of a nitrogen molecule (molar mass 28.0 g/mol) equal that required to escape? What about a hydrogen molecule (molar mass 2.02 g/mol)? (c) Repeat part (b) for the moon, for which $g = 1.63$ m/s^2 and $R_p = 1740$ km. (d) While the earth and the moon have similar average surface temperatures, the moon has essentially no atmosphere. Use your results from parts (b) and (c) to explain why.

18.72 •• Helium gas is in a cylinder that has rigid walls. If the pressure of the gas is 2.00 atm, then the root-mean-square speed of the helium atoms is $v_{rms} = 176$ m/s. By how much (in atmospheres) must the pressure be increased to increase the v_{rms} of the He atoms by 100 m/s? Ignore any change in the volume of the cylinder.

18.73 •• CALC Calculate the integral in Eq. (18.31), $\int_0^\infty v^2 f(v)\,dv$, and compare this result to $(v^2)_{av}$ as given by Eq. (18.16). (*Hint:* You may use the tabulated integral

$$\int_0^\infty x^{2n} e^{-\alpha x^2} dx = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2^{n+1}\alpha^n}\sqrt{\frac{\pi}{\alpha}}$$

where n is a positive integer and α is a positive constant.)

18.74 •• (a) Calculate the total *rotational* kinetic energy of the molecules in 1.00 mol of a diatomic gas at 300 K. (b) Calculate the moment of inertia of an oxygen molecule (O_2) for rotation about either the y- or z-axis shown in Fig. 18.18b. Treat the molecule as two massive points (representing the oxygen atoms) separated by a distance of 1.21×10^{-10} m. The molar mass of oxygen *atoms* is 16.0 g/mol. (c) Find the rms angular velocity of rotation of an oxygen molecule about either the y- or z-axis shown in Fig. 18.18b. How does your answer compare to the angular velocity of a typical piece of rapidly rotating machinery (10,000 rev/min)?

18.75 •• CALC (a) Explain why in a gas of N molecules, the number of molecules having speeds in the *finite* interval v to $v + \Delta v$ is $\Delta N = N\int_v^{v+\Delta v} f(v)\,dv$. (b) If Δv is small, then $f(v)$ is approximately constant over the interval and $\Delta N \approx Nf(v)\Delta v$. For oxygen gas ($O_2$, molar mass 32.0 g/mol) at $T = 300$ K, use this approximation to calculate the number of molecules with speeds within $\Delta v = 20$ m/s of v_{mp}. Express your answer as a multiple of N. (c) Repeat part (b) for speeds within $\Delta v = 20$ m/s of $7v_{mp}$. (d) Repeat parts (b) and (c) for a temperature of 600 K. (e) Repeat parts (b) and (c) for a temperature of 150 K. (f) What do your results tell you about the shape of the distribution as a function of temperature? Do your conclusions agree with what is shown in Fig. 18.23?

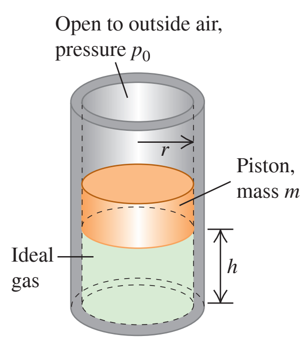
18.76 •• CALC Calculate the integral in Eq. (18.30), $\int_0^\infty v f(v)\,dv$, and compare this result to v_{av} as given by Eq. (18.35). (*Hint:* Make the change of variable $v^2 = x$ and use the tabulated integral

$$\int_0^\infty x^n e^{-\alpha x} dx = \frac{n!}{\alpha^{n+1}}$$

where n is a positive integer and α is a positive constant.)

18.77 ••• CP Oscillations of a Piston. A vertical cylinder of radius r contains an ideal gas and is fitted with a piston of mass m that is free to move (**Fig. P18.77**). The piston and the walls of the cylinder are frictionless, and the entire cylinder is placed in a constant-temperature bath. The outside air pressure is p_0. In equilibrium, the piston sits at a height h above the bottom of the cylinder. (a) Find the absolute pressure of the gas trapped below the piston when in equilibrium. (b) The piston is pulled up by a small distance and released. Find the net force acting on the piston when its base is a distance $h + y$ above the bottom of the cylinder, where $y \ll h$. (c) After the piston is displaced from equilibrium and released, it oscillates up and down. Find the frequency of these small oscillations. If the displacement is not small, are the oscillations simple harmonic? How can you tell?

Figure **P18.77**



18.78 •• DATA A steel cylinder with rigid walls is evacuated to a high degree of vacuum; you then put a small amount of helium into the cylinder. The cylinder has a pressure gauge that measures the pressure of the gas inside the cylinder. You place the cylinder in various temperature environments, wait for thermal equilibrium to be established, and then measure the pressure of the gas. You obtain these results:

	T (°C)	*p* (Pa)
Normal boiling point of nitrogen	−195.8	254
Ice–water mixture	0.0	890
Outdoors on a warm day	33.3	999
Normal boiling point of water	100.0	1214
Hot oven	232	1635

(a) Recall (Chapter 17) that absolute zero is the temperature at which the pressure of an ideal gas becomes zero. Use the data in the table to calculate the value of absolute zero in °C. Assume that the pressure of the gas is low enough for it to be treated as an ideal gas, and ignore the change in volume of the cylinder as its temperature is changed. (b) Use the coefficient of volume expansion for steel in Table 17.2 to calculate the percentage change in the volume of the cylinder between the lowest and highest temperatures in the table. Is it accurate to ignore the volume change of the cylinder as the temperature changes? Justify your answer.

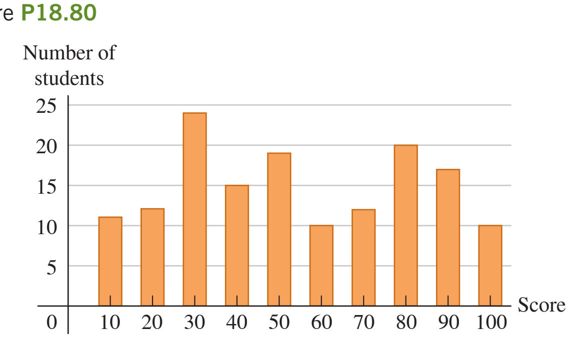
18.79 ••• DATA The Dew Point and Clouds. The vapor pressure of water (see Exercise 18.46) decreases as the temperature decreases. The table lists the vapor pressure of water at various temperatures:

Temperature (°C)	Vapor Pressure (Pa)
10.0	1.23×10^3
12.0	1.40×10^3
14.0	1.60×10^3
16.0	1.81×10^3
18.0	2.06×10^3
20.0	2.34×10^3
22.0	2.65×10^3
24.0	2.99×10^3
26.0	3.36×10^3
28.0	3.78×10^3
30.0	4.25×10^3

If the amount of water vapor in the air is kept constant as the air is cooled, the *dew point* temperature is reached, at which the partial pressure and vapor pressure coincide and the vapor is saturated. If the air is cooled further, the vapor condenses to liquid until the partial pressure again equals the vapor pressure at that temperature. The temperature in a room is 30.0°C. (a) A meteorologist cools a metal can by gradually adding cold water. When the can's temperature reaches 16.0°C, water droplets form on its outside surface. What is the relative humidity of the 30.0°C air in the room? On a spring day in the midwestern United States, the air temperature at the surface is 28.0°C. Puffy cumulus clouds form at an altitude where the air temperature equals the dew point. If the air temperature decreases with altitude at a rate of 0.6 C°/100 m, at approximately what height above the ground will clouds form if the relative humidity at the surface is (b) 35%; (c) 80%?

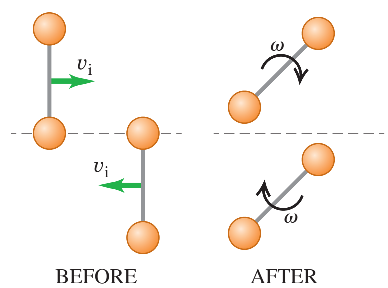
18.80 •• DATA The statistical quantities "average value" and "root-mean-square value" can be applied to any distribution. **Figure P18.80** shows the scores of a class of 150 students on a 100 point quiz. (a) Find the average score for the class. (b) Find the rms score for the class. (c) Which is higher: the average score or the rms score? Why?

Figure **P18.80**



18.81 ••• CP We can crudely model a nitrogen molecule as a pair of small balls, each with the mass of a nitrogen atom, 2.3×10^{-26} kg, attached by a rigid massless rod with length $d = 2r = 188$ pm. (a) What is the moment of inertia of this molecule with respect to an axis passing through the midpoint and perpendicular to the molecular axis? (b) Consider air at 1 atm pressure and 20.0°C temperature. Suppose two nitrogen molecules with rms speeds v_i have an elastic collision such that only one nitrogen atom in one molecule collides with one nitrogen atom in the other molecule in the manner depicted in **Fig. P18.81**. Write equations for the conservation of energy and the conservation of angular momentum, in terms of r, v_i, v_f, and ω, where v_f and ω are, respectively, the center-of-mass speed and the angular speed of either molecule after the collision. (c) Solve these equations for v_f and ω in terms of v_i and r. Note that $\omega \neq 0$ after the collision. (d) Using $v_i = v_{rms}$ for the conditions specified above, what is the frequency of rotation?

Figure **P18.81**



CHALLENGE PROBLEMS

18.82 ••• A pneumatic lift consists of a vertical cylinder with a radius of 10.0 cm. A movable piston slides within the cylinder at its upper end and supports a platform on which loads are placed. An intake valve allows compressed air from a tank to enter the cylinder, and an exhaust valve allows air to be removed from the cylinder. In either case the rate of air transfer is sufficiently low that the temperature in the cylinder remains constant. When neither valve is activated, the cylinder is airtight. The piston and platform together have a mass of 50.0 kg, the temperature is 20.0°C, and the pressure outside the cylinder is 1.00 atm. (a) There is 1.00 mol of air in the cylinder and no load on the platform. What is the height h between the bottom of the piston and the bottom of the cylinder? (b) A 200 kg load is placed on the platform. By what distance does the platform drop? (c) The intake valve is activated and compressed air enters the cylinder so that the platform moves back to its original height. How many moles of air were introduced? (d) How many more moles of air should be introduced so that the platform rises 2.00 m above its original height? (e) At what rate should this air be introduced, in mmol/s, so that the platform rises at a rate of 10.0 cm/s?

18.83 ••• CP Dark Nebulae and the Interstellar Medium. The dark area in **Fig. P18.83** that appears devoid of stars is a *dark nebula,* a cold gas cloud in interstellar space that contains enough material to block out light from the stars behind it. A typical dark nebula is about 20 light-years in diameter and contains about 50 hydrogen atoms per cubic centimeter (monatomic hydrogen, *not* H_2) at about 20 K. (A light-year is the distance light travels in vacuum in one year and is equal to 9.46×10^{15} m.) (a) Estimate the mean free path for a hydrogen atom in a dark nebula. The radius of a hydrogen atom is 5.0×10^{-11} m. (b) Estimate the rms speed of a hydrogen atom and the mean free time (the average time between collisions for a given atom). Based on this result, do you think that atomic collisions, such as those leading to H_2 molecule formation, are very important in determining the composition of the nebula? (c) Estimate the pressure inside a dark nebula. (d) Compare the rms speed of a hydrogen atom to the escape speed at the surface of the nebula (assumed spherical). If the space around the nebula were a vacuum, would such a cloud be stable or would it tend to evaporate? (e) The stability of dark nebulae is explained by the presence of the *interstellar medium* (ISM), an even thinner gas that permeates space and in which the dark nebulae are embedded. Show that for dark nebulae to be in equilibrium with the ISM, the numbers of atoms per volume (N/V) and the temperatures (T) of dark nebulae and the ISM must be related by

$$\frac{(N/V)_{\text{nebula}}}{(N/V)_{\text{ISM}}} = \frac{T_{\text{ISM}}}{T_{\text{nebula}}}$$

(f) In the vicinity of the sun, the ISM contains about 1 hydrogen atom per 200 cm^3. Estimate the temperature of the ISM in the vicinity of the sun. Compare to the temperature of the sun's surface, about 5800 K. Would a spacecraft coasting through interstellar space burn up? Why or why not?

Figure **P18.83**

18.84 ••• CALC Earth's Atmosphere. In the *troposphere,* the part of the atmosphere that extends from earth's surface to an altitude of about 11 km, the temperature is not uniform but decreases with increasing elevation. (a) Show that if the temperature variation is approximated by the linear relationship

$$T = T_0 - \alpha y$$

where T_0 is the temperature at the earth's surface and T is the temperature at height y, the pressure p at height y is

$$\ln\left(\frac{p}{p_0}\right) = \frac{Mg}{R\alpha}\ln\left(\frac{T_0 - \alpha y}{T_0}\right)$$

where p_0 is the pressure at the earth's surface and M is the molar mass for air. The coefficient α is called the lapse rate of temperature. It varies with atmospheric conditions, but an average value is about 0.6 C°/100 m. (b) Show that the above result reduces to the result of Example 18.4 (Section 18.1) in the limit that $\alpha \rightarrow 0$. (c) With $\alpha = 0.6$ C°/100 m, calculate p for $y = 8863$ m and compare your answer to the result of Example 18.4. Take $T_0 = 288$ K and $p_0 = 1.00$ atm.

MCAT-STYLE PASSAGE PROBLEMS

Insulating Windows. One way to improve insulation in windows is to fill a sealed space between two glass panes with a gas that has a lower thermal conductivity than that of air. The thermal conductivity k of a gas depends on its molar heat capacity C_V, molar mass M, and molecular radius r. The dependence on those quantities at a given temperature is approximated by $k \propto C_V/r^2\sqrt{M}$. The noble gases have properties that make them particularly good choices as insulating gases. Noble gases range from helium (molar mass 4.0 g/mol, molecular radius 0.13 nm) to xenon (molar mass 131 g/mol, molecular radius 0.22 nm). (The noble gas radon is heavier than xenon, but radon is radioactive and so is not suitable for this purpose.)

18.85 What is one reason the noble gases are *preferable* to air (which is mostly nitrogen and oxygen) as an insulating material? (a) Noble gases are monatomic, so no rotational modes contribute to their molar heat capacity; (b) noble gases are monatomic, so they have lower molecular masses than do nitrogen and oxygen; (c) molecular radii in noble gases are much larger than those of gases that consist of diatomic molecules; (d) because noble gases are monatomic, they have many more degrees of freedom than do diatomic molecules, and their molar heat capacity is reduced by the number of degrees of freedom.

18.86 Estimate the ratio of the thermal conductivity of Xe to that of He. (a) 0.015; (b) 0.061; (c) 0.10; (d) 0.17.

18.87 The rate of *effusion*—that is, leakage of a gas through tiny cracks—is proportional to v_{rms}. If tiny cracks exist in the material that's used to seal the space between two glass panes, how many times greater is the rate of He leakage out of the space between the panes than the rate of Xe leakage at the same temperature? (a) 370 times; (b) 19 times; (c) 6 times; (d) no greater—the He leakage rate is the same as for Xe.

ANSWERS

Chapter Opening Question ?

(iv) From Eq. (18.19), the root-mean-square speed of a gas molecule is proportional to the square root of the absolute temperature T. The temperature range we're considering is $(25 + 273.15)\text{ K} = 298\text{ K}$ to $(100 + 273.15)\text{ K} = 373\text{ K}$. Hence the speeds increase by a factor of $\sqrt{(373\text{ K})/(298\text{ K})} = 1.12$; that is, there is a 12% increase. While 100°C feels far warmer than 25°C, the difference in molecular speeds is relatively small.

Key Example VARIATION Problems

VP18.4.1 (a) 4.67×10^5 Pa (b) 0.278 mol

VP18.4.2 (a) $1.20 \times 10^3\text{ m}^3$, (b) 4.52 times

VP18.4.3 (a) $2.9 \times 10^{-3}\text{ mol/m}^3$ (b) $8.0 \times 10^{-5}\text{ kg/m}^3$

VP18.4.4 (a) $T_2 = T_1\left(\frac{p_2}{p_1}\right)^{2/5}$ (b) $\frac{\rho_2}{\rho_1} = 0.660$, $\frac{T_2}{T_1} = 0.758$ (c) $\frac{\rho_2}{\rho_1} - 1.52$, $\frac{T_2}{T_1} = 1.32$

VP18.7.1 −231°C

VP18.7.2 (a) 6.2×10^{-20} J (b) 8.6×10^3 m/s

VP18.7.3 (a) 1.5×10^{27} (b) 9.1×10^6 J (c) 1.1×10^2 m/s, or 4.0×10^2 km/h

VP18.7.4 (a) 5.50 (b) 38.5 (c) 6.20

VP18.8.1 (a) 6.8×10^{-6} m (b) 1.2×10^2 times greater

VP18.8.2 (a) 5.7×10^{-3} Pa (b) 2.3×10^{-6} mol

VP18.8.3 (a) 1.2×10^{-8} m (b) 8.9×10^{-12}

VP18.8.4 (a) $t = \frac{1}{4\pi r^2 p}\sqrt{\frac{mkT}{6}}$ (b) doubling r

Bridging Problem

(a) 102 km

(b) yes

? A steam locomotive uses the first law of thermodynamics: Water is heated and boils, and the expanding steam does work to propel the locomotive. Would it be possible for the steam to propel the locomotive by doing work as it *condenses?* (i) Yes; (ii) no; (iii) answer depends on the details of how the steam condenses.

19 The First Law of Thermodynamics

Every time you drive a gasoline-powered car, turn on an air conditioner, or cook a meal, you reap the benefits of *thermodynamics,* the study of relationships involving heat, mechanical work, and other aspects of energy and energy transfer. For example, in a car engine heat is generated by the chemical reaction of oxygen and vaporized gasoline in the engine's cylinders. The heated gas pushes on the pistons within the cylinders, doing mechanical work that is used to propel the car. This is an example of a *thermodynamic process.*

The first law of thermodynamics, central to the understanding of such processes, is an extension of the principle of conservation of energy. It broadens this principle to include energy exchange by both heat transfer and mechanical work and introduces the concept of the *internal energy* of a system. Conservation of energy plays a vital role in every area of physical science, and the first law has extremely broad usefulness. To state energy relationships precisely, we need the concept of a *thermodynamic system.* We'll discuss *heat* and *work* as two means of transferring energy into or out of such a system.

LEARNING OUTCOMES

In this chapter, you'll learn...

- **19.1** The significance of thermodynamic systems and processes.
- **19.2** How to calculate work done by a system when its volume changes.
- **19.3** What is meant by a path between thermodynamic states.
- **19.4** How to interpret and use the first law of thermodynamics.
- **19.5** Four important kinds of thermodynamic processes.
- **19.6** Why the internal energy of an ideal gas depends on temperature only.
- **19.7** The difference between molar heat capacities at constant volume and at constant pressure.
- **19.8** How to analyze adiabatic processes in an ideal gas.

You'll need to review...

- **6.3** Work done by a force.
- **7.3** Internal energy.
- **17.5, 18.4** Specific heat and molar heat capacity.
- **18.1** *pV*-diagrams.

19.1 THERMODYNAMIC SYSTEMS

We have studied energy transfer through mechanical work (Chapter 6) and through heat transfer (Chapters 17 and 18). Now we are ready to combine and generalize these principles.

We always talk about energy transfer to or from some specific *system.* The system might be a mechanical device, a biological organism, or a specified quantity of material, such as the refrigerant in an air conditioner or steam expanding in a turbine. In general, a **thermodynamic system** is any collection of objects that is convenient to regard as a unit, and that may have the potential to exchange energy with its surroundings. A familiar example is a quantity of popcorn kernels in a pot with a lid. When the pot is placed on a stove, energy is added to the popcorn by conduction of heat. As the popcorn pops and expands, it does work as it exerts an upward force on the lid and moves it through a displacement (**Fig. 19.1**). The *state* of the popcorn—its volume, temperature,

Figure 19.1 The popcorn in the pot is a thermodynamic system. In the thermodynamic process shown here, heat is added to the system, and the system does work on its surroundings to lift the lid of the pot.

and pressure—changes as it pops. A process such as this one, in which there are changes in the state of a thermodynamic system, is called a **thermodynamic process.**

In mechanics we used the concept of *system* with free-body diagrams and with conservation of energy and momentum. For *thermodynamic* systems, as for all others, it is essential to define clearly at the start exactly what is and is not included in the system. Only then can we describe unambiguously the energy transfers into and out of that system. For instance, in our popcorn example we defined the system to include the popcorn but not the pot, lid, or stove.

Thermodynamics has its roots in many practical problems other than popping popcorn (**Fig. 19.2**). The gasoline engine in an automobile, the jet engines in an airplane, and the rocket engines in a launch vehicle use the heat of combustion of their fuel to perform mechanical work in propelling the vehicle. Muscle tissue in living organisms metabolizes chemical energy in food and performs mechanical work on the organism's surroundings. A steam engine or steam turbine uses the heat of combustion of coal or other fuel to perform mechanical work such as driving an electric generator or pulling a train.

Signs for Heat and Work in Thermodynamics

We describe the energy relationships in any thermodynamic process in terms of the quantity of heat Q added *to* the system and the work W done *by* the system. Both Q and W may be positive, negative, or zero (**Fig. 19.3**). A positive value of Q represents heat flow *into* the system, with a corresponding input of energy to it; negative Q represents heat flow *out of* the system. A positive value of W represents work done *by* the system against its surroundings, such as work done by an expanding gas, and hence corresponds to energy *leaving* the system. Negative W, such as work done during compression of a gas in which work is done *on the gas* by its surroundings, represents energy *entering* the system. We'll use these conventions consistently in the examples in this chapter and the next.

CAUTION Be careful with the sign of work W Note that our sign rule for work is *opposite* to the one we used in mechanics, in which we always spoke of the work done by the forces acting *on* an object. In thermodynamics it is usually more convenient to call W the work done *by* the system so that when a system expands, the pressure, volume change, and work are all positive. Use the sign rules for work and heat consistently!

TEST YOUR UNDERSTANDING OF SECTION 19.1 In Example 17.7 (Section 17.6), what is the sign of Q for the coffee? For the aluminum cup? If a block slides along a horizontal surface with friction, what is the sign of W for the block?

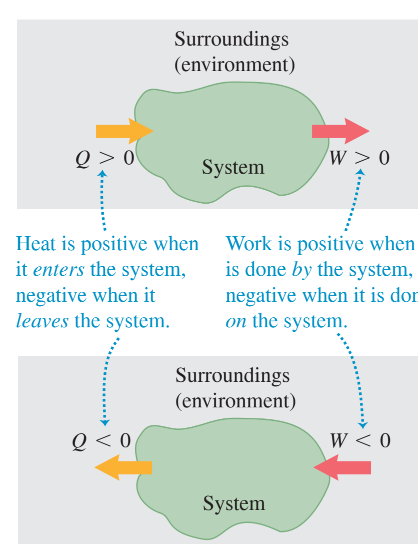
ANSWER

negative, positive, positive Heat flows out of the coffee, so $Q_{\text{coffee}} < 0$; heat flows into the aluminum cup, so $Q_{\text{aluminum}} > 0$. In mechanics, we would say that negative work is done *on* the block, since the surface exerts a force on the block that opposes the block's motion. But in thermodynamics we use the opposite convention and say that $W > 0$, which means that positive work is done *by* the block on the surface.

Figure 19.2 (a) A rocket engine uses the heat of combustion of its fuel to do work propelling the launch vehicle. (b) Humans and other biological organisms are more complicated systems than we can analyze fully in this book, but the same basic principles of thermodynamics apply to them.

(a) (b)

Figure 19.3 A thermodynamic system may exchange energy with its surroundings (environment) by means of heat, work, or both. Note the sign conventions for Q and W.



19.2 WORK DONE DURING VOLUME CHANGES

A simple example of a thermodynamic system is a quantity of gas enclosed in a cylinder with a movable piston. Internal-combustion engines, steam engines, and compressors in refrigerators and air conditioners all use some version of such a system. In the next several sections we'll use the gas-in-cylinder system to explore several kinds of thermodynamic processes.

We'll use a microscopic viewpoint, based on the kinetic and potential energies of individual molecules in a material, to develop intuition about thermodynamic quantities. But it is important to understand that the central principles of thermodynamics can be treated in a completely *macroscopic* way, without reference to microscopic models. Indeed, part of the great power and generality of thermodynamics is that it does *not* depend on details of the structure of matter.

First we consider the *work* done by the system during a volume change. When a gas expands, it pushes outward on its boundary surfaces as they move outward. Hence an expanding gas always does positive work. The same thing is true of any material that expands under pressure, such as the popcorn in Fig. 19.1.

We can understand the work done by a gas in a volume change by considering the molecules that make up the gas. When one such molecule collides with a stationary surface, it exerts a momentary force on the wall but does no work because the wall does not move. But if the surface is moving, like a piston in a gasoline engine, the molecule *does* do work on the surface during the collision. If the piston in **Fig. 19.4a** moves to the right, so the volume of the gas increases, the molecules that strike the piston exert a force through a distance and do *positive* work on the piston. If the piston moves toward the left as in Fig. 19.4b, so the volume of the gas decreases, positive work is done *on* the molecule during the collision. Hence the gas molecules do *negative* work on the piston.

Figure 19.4 A molecule striking a piston (a) does positive work if the piston is moving away from the molecule and (b) does negative work if the piston is moving toward the molecule. Hence a gas does positive work when it expands as in (a) but does negative work when it compresses as in (b).

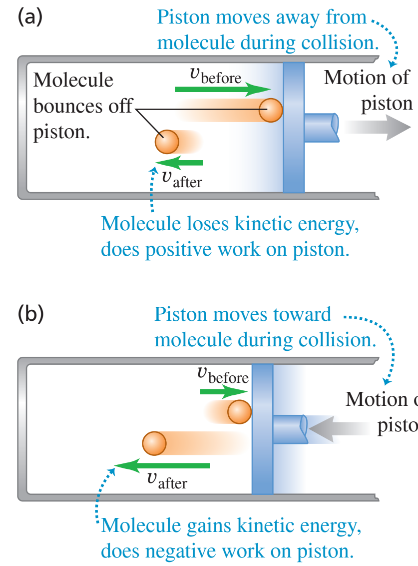


Figure 19.5 shows a system whose volume can change (a gas, liquid, or solid) in a cylinder with a movable piston. Suppose that the cylinder has cross-sectional area A and that the pressure exerted by the system at the piston face is p. The total force F exerted by the system on the piston is $F = pA$. When the piston moves out an infinitesimal distance dx, the work dW done by this force is

$$dW = F\,dx = pA\,dx$$

But

$$A\,dx = dV$$

where dV is the infinitesimal change of volume of the system. Thus we can express the work done by the system in this infinitesimal volume change as

$$dW = p\,dV \tag{19.1}$$

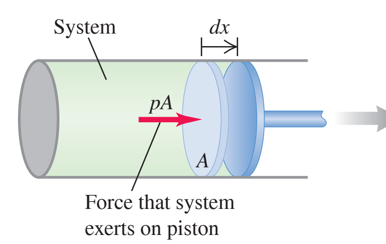
In a finite change of volume from V_1 to V_2,

Work done in a volume change

$$W = \int_{V_1}^{V_2} p\,dV \tag{19.2}$$

Upper limit = final volume
Integral of the pressure with respect to volume
Lower limit = initial volume

Figure 19.5 The infinitesimal work done by the system during the small expansion dx is $dW = pA\,dx$.



In general, the pressure of the system may vary during the volume change. For example, this is the case in the cylinders of an automobile engine as the pistons move back and forth. To evaluate the integral in Eq. (19.2), we have to know how the pressure varies as a function of volume. We can represent this relationship as a graph of p as a function of V (a pV-diagram, described at the end of Section 18.1). **Figure 19.6** (next page) shows a simple example. In this figure, Eq. (19.2) is represented graphically as the *area* under the curve of p versus V between the limits V_1 and V_2. (In Section 6.3 we used a similar interpretation of the work done by a force F as the area under the curve of F versus x between the limits x_1 and x_2.)

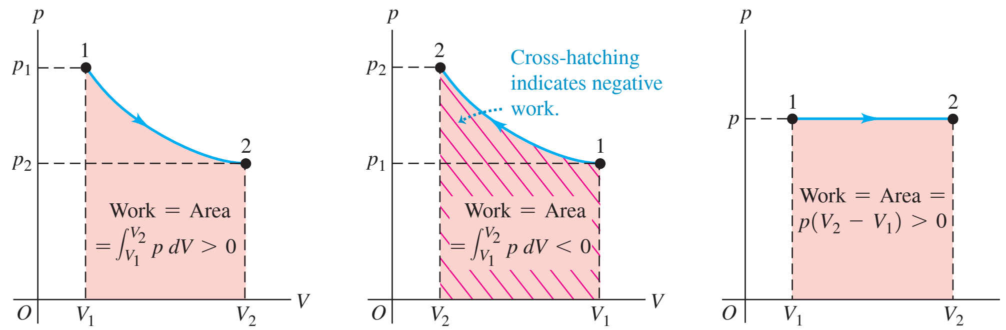
According to the rule we stated in Section 19.1, work is *positive* when a system *expands*. In an expansion from state 1 to state 2 in Fig. 19.6a, the area under the curve and

Figure **19.6** The work done equals the area under the curve on a pV-diagram.

(a) pV-diagram for a system undergoing an expansion with varying pressure

(b) pV-diagram for a system undergoing a compression with varying pressure

(c) pV-diagram for a system undergoing an expansion with constant pressure



the work are positive. A *compression* from 1 to 2 in Fig. 19.6b gives a *negative* area; when a system is compressed, its volume decreases and it does *negative* work on its surroundings (see also Fig. 19.4b).

If the pressure p remains constant while the volume changes from V_1 to V_2 (Fig. 19.6c), the work done by the system is

Work done in a volume change at constant pressure

Pressure

$$W = p(V_2 - V_1) \qquad (19.3)$$

Final volume Initial volume

If the volume is *constant,* there is no displacement and the system does no work.

CAUTION Be careful with subscripts 1 and 2 In Eq. (19.2), V_1 is the *initial* volume and V_2 is the *final* volume. That's why labels 1 and 2 are reversed in Fig. 19.6b compared to Fig. 19.6a, even though both processes move between the same two thermodynamic states.

EXAMPLE 19.1 Isothermal expansion of an ideal gas

As an ideal gas undergoes an *isothermal* (constant-temperature) expansion at temperature T, its volume changes from V_1 to V_2. How much work does the gas do?

IDENTIFY and SET UP The ideal-gas equation, Eq. (18.3), tells us that if the temperature T of n moles of an ideal gas is constant, the quantity $pV = nRT$ is also constant. If V changes, p changes as well, so we *cannot* use Eq. (19.3) to calculate the work done. Instead we must evaluate the integral in Eq. (19.2), so we must know p as a function of V; for this we use Eq. (18.3).

EXECUTE From Eq. (18.3),

$$p = \frac{nRT}{V}$$

We substitute this into the integral of Eq. (19.2), take the constant factor nRT outside, and evaluate the integral:

$$W = \int_{V_1}^{V_2} p\,dV = nRT \int_{V_1}^{V_2} \frac{dV}{V}$$

$$= nRT \ln \frac{V_2}{V_1} \qquad \text{(ideal gas, isothermal process)}$$

We can rewrite this expression for W in terms of p_1 and p_2. Because $pV = nRT$ is constant,

$$p_1V_1 = p_2V_2 \qquad \text{or} \qquad \frac{V_2}{V_1} = \frac{p_1}{p_2}$$

so

$$W = nRT \ln \frac{p_1}{p_2} \qquad \text{(ideal gas, isothermal process)}$$

EVALUATE We check our result by noting that in an expansion $V_2 > V_1$ and the ratio V_2/V_1 is greater than 1. The logarithm of a number greater than 1 is positive, so $W > 0$, as it should be. As an additional check, look at our second expression for W: In an isothermal expansion the volume increases and the pressure drops, so $p_2 < p_1$, the ratio $p_1/p_2 > 1$, and $W = nRT \ln(p_1/p_2)$ is again positive.

Our result for W also applies to an isothermal *compression* of a gas, for which $V_2 < V_1$ and $p_2 > p_1$.

KEYCONCEPT To find the amount of work a system does on its surroundings during a volume change, you can either calculate the integral of pressure p with respect to volume V, or find the area under the curve of p versus V. The work is positive when the volume increases, and negative when the volume decreases.

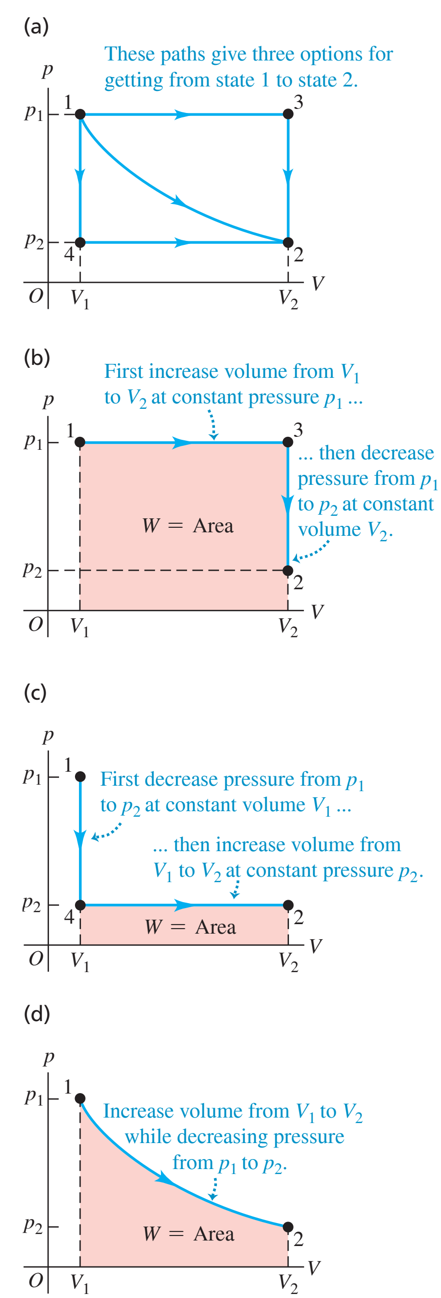
TEST YOUR UNDERSTANDING OF SECTION 19.2 A quantity of ideal gas undergoes an expansion that doubles its volume. Does the gas do more work on its surroundings if the expansion is at constant *pressure* or at constant *temperature*? (i) Constant pressure; (ii) constant temperature; (iii) the same amount in both cases; (iv) not enough information is given to decide.

ANSWER

(ii) The work done in an expansion is represented by the area under the curve of pressure p versus volume V. In an isothermal expansion the pressure decreases as the volume increases, so the pV-diagram looks like Fig. 19.6a and the work done equals the shaded area under the blue curve from point 1 to point 2. If, however, the expansion is at constant pressure, the curve of p versus V would be the same as the dashed horizontal line at pressure p_2 in Fig. 19.6a. The area under this dashed line is smaller than the area under the blue curve for an isothermal expansion, so less work is done in the constant-pressure expansion than in the isothermal expansion.

19.3 PATHS BETWEEN THERMODYNAMIC STATES

Figure 19.7 The work done by a system during a transition between two states depends on the path chosen.



We've seen that if a thermodynamic process involves a change in volume, the system undergoing the process does work (either positive or negative) on its surroundings. Heat also flows into or out of the system during the process if there is a temperature difference between the system and its surroundings. Let's now examine how the work done by and the heat added to the system during a thermodynamic process depend on the details of how the process takes place.

Work Done in a Thermodynamic Process

When a thermodynamic system changes from an initial state to a final state, it passes through a series of intermediate states. We call this series of states a **path.** There are always infinitely many possibilities for these intermediate states. When all are equilibrium states, the path can be plotted on a pV-diagram (**Fig. 19.7a**). Point 1 represents an initial state with pressure p_1 and volume V_1, and point 2 represents a final state with pressure p_2 and volume V_2. To pass from state 1 to state 2, we could keep the pressure constant at p_1 while the system expands to volume V_2 (point 3 in Fig. 19.7b), then reduce the pressure to p_2 (probably by decreasing the temperature) while keeping the volume constant at V_2 (to point 2). The work done by the system during this process is the area under the line $1 \longrightarrow 3$; no work is done during the constant-volume process $3 \longrightarrow 2$. Or the system might traverse the path $1 \longrightarrow 4 \longrightarrow 2$ (Fig. 19.7c); then the work is the area under the line $4 \longrightarrow 2$, since no work is done during the constant-volume process $1 \longrightarrow 4$. The smooth curve from 1 to 2 is another possibility (Fig. 19.7d), and the work for this path is different from that for either of the other paths.

We conclude that *the work done by the system depends not only on the initial and final states, but also on the intermediate states—that is, on the path.* Furthermore, we can take the system through a series of states forming a closed loop, such as $1 \longrightarrow 3 \longrightarrow 2 \longrightarrow 4 \longrightarrow 1$. In this case the final state is the same as the initial state, but the total work done by the system is *not* zero. (In fact, it is represented on the graph by the area enclosed by the loop; see Exercise 19.7.) So we can't talk about the amount of work *contained in* a system. In a particular state, a system may have definite values of the state coordinates p, V, and T, but it wouldn't make sense to say that it has a definite value of W.

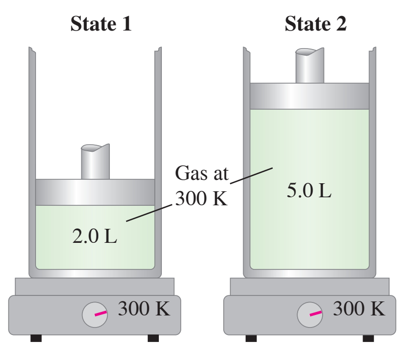
Heat Added in a Thermodynamic Process

Like work, the *heat* added to a thermodynamic system when it undergoes a change of state depends on the path from the initial state to the final state. Here's an example. Suppose we want to change the volume of a certain quantity of an ideal gas from 2.0 L to 5.0 L while keeping the temperature constant at $T = 300$ K. **Figure 19.8** (next page) shows two different ways to do this. In Fig. 19.8a the gas is contained in a cylinder with a piston, with an initial volume of 2.0 L. We let the gas expand slowly, supplying heat from the electric heater to keep the temperature at 300 K until the gas reaches its final volume of 5.0 L. The gas absorbs a definite amount of heat in this isothermal process.

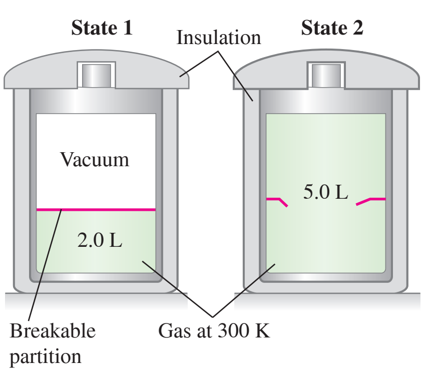
Figure 19.8b shows a different process leading to the same final state. The container is surrounded by insulating walls and is divided by a thin, breakable partition into two compartments. The lower part has volume 2.0 L and the upper part has volume 3.0 L. In the lower compartment we place the same amount of the same gas as in Fig. 19.8a,

Figure 19.8 (a) Slow, controlled isothermal expansion of a gas from an initial state 1 to a final state 2 with the same temperature but lower pressure. (b) Rapid, uncontrolled expansion of the same gas starting at the same state 1 and ending at the same state 2.

(a) System does work on piston; hot plate adds heat to system ($W > 0$ and $Q > 0$).



(b) System does no work; no heat enters or leaves system ($W = 0$ and $Q = 0$).



again at $T = 300$ K. The initial state is the same as before. Now we break the partition; the gas expands rapidly, with no heat passing through the insulating walls. The final volume is 5.0 L, the same as in Fig. 19.8a. The expanding gas does no work because it doesn't push against anything that moves. This uncontrolled expansion of a gas into vacuum is called a **free expansion;** we'll discuss it further in Section 19.6.

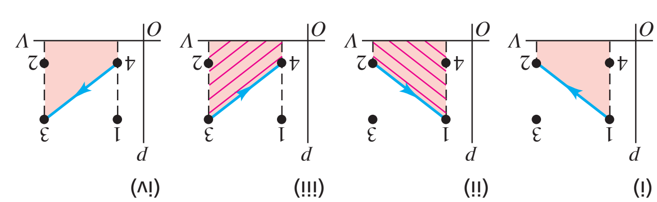
Experiments show that when an ideal gas undergoes a free expansion, there is no temperature change. So the final state of the gas is the same as in Fig. 19.8a. The intermediate pressures and volumes during the transition from state 1 to state 2 are entirely different in the two cases; Figs. 19.8a and 19.8b represent *two different paths* connecting the *same states* 1 and 2. For the path in Fig. 19.8b, *no* heat is transferred into the system, and the system does no work. Like work, *heat depends not only on the initial and final states but also on the path.*

Because of this path dependence, it would not make sense to say that a system "contains" a certain quantity of heat. To see this, suppose we assign an arbitrary value to the "heat in an object" in some reference state. Then presumably the "heat in the object" in some other state would equal the heat in the reference state plus the heat added when the object goes to the second state. But that's ambiguous, as we have just seen; the heat added depends on the *path* we take from the reference state to the second state. We are forced to conclude that there is *no* consistent way to define "heat in an object"; it is not a useful concept.

While it doesn't make sense to talk about "work in an object" or "heat in an object," it *does* make sense to speak of the amount of *internal energy* in an object. This important concept is our next topic.

TEST YOUR UNDERSTANDING OF SECTION 19.3 The system described in Fig. 19.7a undergoes four different thermodynamic processes. Each process is represented in a pV-diagram as a straight line from the initial state to the final state. (These processes are different from those shown in the pV-diagrams of Fig. 19.7.) Rank the processes in order of the amount of work done by the system, from the most positive to the most negative. (i) $1 \rightarrow 2$; (ii) $2 \rightarrow 1$; (iii) $3 \rightarrow 4$; (iv) $4 \rightarrow 3$.

ANSWER



(i) and (iv) (tie), (ii) and (iii) (tie) The accompanying figure shows the pV-diagrams for each of the four processes. The trapezoidal area under the curve, and hence the absolute value of the work, is the same in all four cases. In cases (i) and (iv) the volume increases, so the system does positive work as it expands against its surroundings. In cases (ii) and (iii) the volume decreases, so the system does negative work (shown by cross-hatching) as the surroundings push inward on it.

19.4 INTERNAL ENERGY AND THE FIRST LAW OF THERMODYNAMICS

Internal energy is one of the most important concepts in thermodynamics. In Section 7.3, when we discussed energy changes for an object sliding with friction, we stated that warming an object increased its internal energy and that cooling the object decreased its internal energy. But what *is* internal energy? We can look at it in various ways; let's start with one based on the ideas of mechanics. Matter consists of atoms and molecules, and these are made up of particles having kinetic and potential energies. We *tentatively* define the **internal energy** of a system as the sum of the kinetic energies of all of its constituent particles, plus the sum of all the potential energies of interaction among these particles.

CAUTION Is it internal? Internal energy does *not* include potential energy arising from the interaction between the system and its surroundings. If the system is a glass of water, placing it on a high shelf increases the gravitational potential energy arising from the interaction between the glass and the earth. But this has no effect on the interactions among the water molecules, and so the internal energy of the water does not change.

We use the symbol U for internal energy. (We used this symbol in our study of mechanics to represent potential energy. However, U has a different meaning in thermodynamics.) During a change of state of the system, the internal energy may change from an initial value U_1 to a final value U_2. We denote the change in internal energy as $\Delta U = U_2 - U_1$.

When we add a quantity of heat Q to a system and the system does no work during the process (so $W = 0$), the internal energy increases by an amount equal to Q; that is, $\Delta U = Q$. When a system does work W by expanding against its surroundings and no heat is added during the process, energy leaves the system and the internal energy decreases: W is positive, Q is zero, and $\Delta U = -W$. When *both* heat transfer and work occur, the *total* change in internal energy is

First law of thermodynamics:

$$\overbrace{\Delta U}^{\text{Internal energy change of thermodynamic system}} = \underbrace{Q}_{\text{Heat added to system}} - \underbrace{W}_{\text{Work done by system}} \qquad (19.4)$$

We can rearrange this to the form

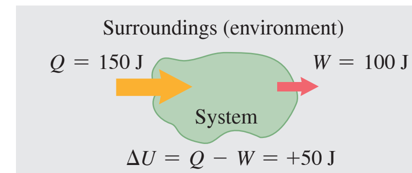
$$Q = \Delta U + W \qquad (19.5)$$

The message of Eq. (19.5) is that when heat Q is added to a system, some of this added energy remains within the system, changing its internal energy by ΔU; the remainder leaves the system as the system does work W on its surroundings. Because W and Q may be positive, negative, or zero, ΔU can be positive, negative, or zero for different processes (**Fig. 19.9**).

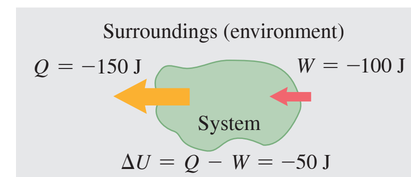
Equation (19.4) or (19.5) is the **first law of thermodynamics.** It is a generalization of the principle of conservation of energy to include energy transfer through heat as well as mechanical work. As you'll see in later chapters, this principle can be extended to ever-broader classes of phenomena by identifying additional forms of energy and energy transfer. In every situation in which it seems that the total energy in all known forms is not conserved, it has been possible to identify a new form of energy such that the total energy, including the new form, *is* conserved.

Figure **19.9** In a thermodynamic process, the internal energy U of a system may (a) increase ($\Delta U > 0$), (b) decrease ($\Delta U < 0$), or (c) remain the same ($\Delta U = 0$).

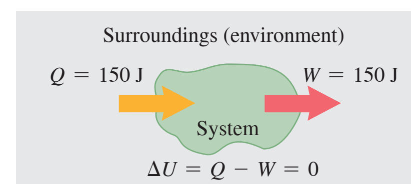
(a) More heat is added to system than system does work: Internal energy of system increases.



(b) More heat flows out of system than work is done: Internal energy of system decreases.



(c) Heat added to system equals work done by system: Internal energy of system unchanged.



Understanding the First Law of Thermodynamics

At the beginning of this discussion we tentatively defined internal energy in terms of microscopic kinetic and potential energies. But actually *calculating* internal energy in this way for any real system would be hopelessly complicated. Furthermore, this definition isn't an *operational* one: It doesn't describe how to determine internal energy from physical quantities that we can measure.

So let's look at internal energy in another way. Starting over, we define the *change* in internal energy ΔU during any change of a system as the quantity given by Eq. (19.4), $\Delta U = Q - W$. This *is* an operational definition because we can measure Q and W. It does not define U itself, only ΔU. This is not a shortcoming because we can *define* the internal energy of a system to have a specified value in some reference state, and then use Eq. (19.4) to define the internal energy in any other state. This is analogous to our treatment of potential energy in Chapter 7, in which we arbitrarily defined the potential energy of a mechanical system to be zero at a certain position.

This new definition trades one difficulty for another. If we define ΔU by Eq. (19.4), then when the system goes from state 1 to state 2 by two different paths, how do we know that ΔU is the same for the two paths? We have already seen that Q and W are, in general, *not* the same for different paths. If ΔU, which equals $Q - W$, is also path dependent, then ΔU is ambiguous. If so, the concept of internal energy of a system is subject to the same criticism as the erroneous concept of quantity of heat in a system, as we discussed at the end of Section 19.3.

The only way to answer this question is through *experiment.* For various materials we measure Q and W for various changes of state and various paths to learn whether ΔU is or is not path dependent. The results of many such investigations are clear and unambiguous: While Q and W depend on the path, $\Delta U = Q - W$ *is independent of path. The change in internal energy of a system during any thermodynamic process depends only on the initial and final states, not on the path leading from one to the other.*

BIO APPLICATION The First Law of Exercise Thermodynamics Your body is a thermodynamic system. When you exercise, your body does work (such as the work done to lift your body as a whole in a push-up). Hence $W > 0$. Your body also warms up during exercise; by perspiration and other means the body rids itself of this heat, so $Q < 0$. Since Q is negative and W is positive, $\Delta U = Q - W < 0$ and the body's internal energy decreases. That's why exercise helps you lose weight: It uses up some of the internal energy stored in your body in the form of fat.

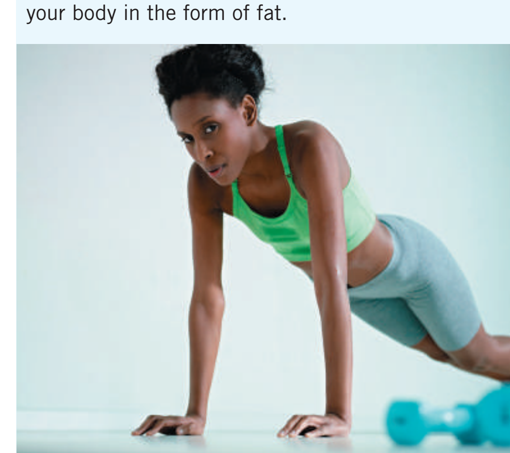


Figure 19.10 The internal energy of a cup of coffee depends on just its thermodynamic state—how much water and ground coffee it contains, and what its temperature is. It does not depend on the history of how the coffee was prepared—that is, the thermodynamic path that led to its current state.

Experiment, then, is the ultimate justification for believing that a thermodynamic system in a specific state has a unique internal energy that depends only on that state. An equivalent statement is that the internal energy U of a system is a function of the state coordinates p, V, and T (actually, any two of these, since the three variables are related by the equation of state).

To say that the first law of thermodynamics, given by Eq. (19.4) or (19.5), represents conservation of energy for thermodynamic processes is correct, as far as it goes. But an important *additional* aspect of the first law is the fact that internal energy depends only on the state of a system (**Fig. 19.10**). In changes of state, the change in internal energy is independent of the path.

All this may seem a little abstract if you are satisfied to think of internal energy as microscopic mechanical energy. There's nothing wrong with that view, and we'll make use of it at various times during our discussion. But as for heat, a precise *operational* definition of internal energy must be independent of the detailed microscopic structure of the material.

Cyclic Processes and Isolated Systems

Two special cases of the first law of thermodynamics are worth mentioning. A process that eventually returns a system to its initial state is called a *cyclic process.* For such a process, the final state is the same as the initial state, and so the *total* internal energy change must be zero. Then

$$U_2 = U_1 \qquad \text{and} \qquad Q = W$$

If a net quantity of work W is done by the system during this process, an equal amount of energy must have flowed into the system as heat Q. But there is no reason either Q or W individually has to be zero (**Fig. 19.11**).

Another special case occurs in an *isolated system,* one that does no work on its surroundings and has no heat flow to or from its surroundings. For any process taking place in an isolated system,

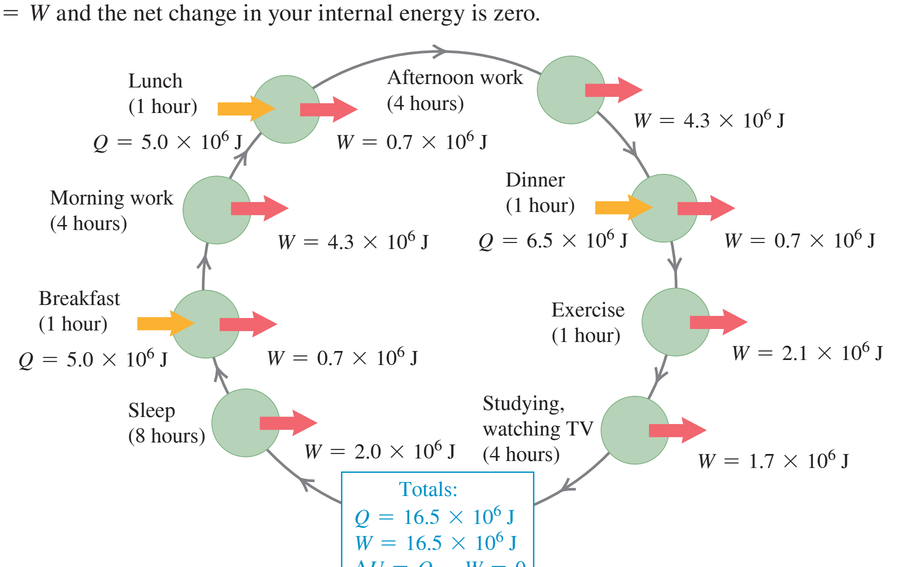
$$W = Q = 0$$

and therefore

$$U_2 = U_1 = \Delta U = 0$$

In other words, *the internal energy of an isolated system is constant.*

Figure 19.11 Every day, your body (a thermodynamic system) goes through a cyclic thermodynamic process like this one. Heat Q is added by metabolizing food, and your body does work W in breathing, walking, and other activities. If you return to the same state at the end of the day, $Q = W$ and the net change in your internal energy is zero.



PROBLEM-SOLVING STRATEGY 19.1 The First Law of Thermodynamics

IDENTIFY *the relevant concepts:* The first law of thermodynamics is the statement of the law of conservation of energy in its most general form. You can apply it to *any* thermodynamic process in which the internal energy of a system changes, heat flows into or out of the system, and/or work is done by or on the system.

SET UP *the problem* using the following steps:

1. Define the thermodynamic system to be considered.
2. If the thermodynamic process has more than one step, identify the initial and final states for each step.
3. List the known and unknown quantities and identify the target variables.
4. Confirm that you have enough equations. You can apply the first law, $\Delta U = Q - W$, just once to each step in a thermodynamic process, so you'll often need additional equations. These may include Eq. (19.2), $W = \int_{V_1}^{V_2} p\,dV$, which gives the work W done in a volume change, and the equation of state of the material that makes up the thermodynamic system (for an ideal gas, $pV = nRT$).

EXECUTE *the solution* as follows:

1. Be sure to use consistent units. If p is in Pa and V in m^3, then W is in joules. If a heat capacity is given in terms of calories, convert it to joules. When you use $n = m_{total}/M$ to relate total mass m_{total} to number of moles n, remember that if m_{total} is in kilograms, M must be in *kilograms* per mole; M is usually tabulated in *grams* per mole.
2. The internal energy change ΔU in any thermodynamic process or series of processes is independent of the path, whether the substance is an ideal gas or not. If you can calculate ΔU for *any* path between given initial and final states, you know ΔU for *every possible path* between those states; you can then relate the various energy quantities for any of those other paths.
3. In a process comprising several steps, tabulate Q, W, and ΔU for each step, with one line per step and with the Q's, W's, and ΔU's forming columns (see Example 19.4). You can apply the first law to each line, and you can add each column and apply the first law to the sums. Do you see why?
4. Using steps 1–3, solve for the target variables.

EVALUATE *your answer:* Check your results for reasonableness. Ensure that each of your answers has the correct algebraic sign. A positive Q means that heat flows *into* the system; a negative Q means that heat flows *out of* the system. A positive W means that work is done *by* the system on its environment; a negative W means that work is done *on* the system by its environment.

EXAMPLE 19.2 Working off your dessert

WITH VARIATION PROBLEMS

You propose to climb several flights of stairs to work off the energy you took in by eating a 900 calorie hot fudge sundae. How high must you climb? Assume that your mass is 60.0 kg.

IDENTIFY and SET UP The thermodynamic system is your body. You climb the stairs to make the final state of the system the same as the initial state (no fatter, no leaner). There is therefore no net change in internal energy: $\Delta U = 0$. Eating the hot fudge sundae corresponds to a heat flow into your body, and you do work climbing the stairs. We can relate these quantities by using the first law of thermodynamics. We are given that $Q = 900$ food calories (900 kcal) of heat flow into your body. The work you must do to raise your mass m a height h is $W = mgh$; our target variable is h.

EXECUTE From the first law of thermodynamics, $\Delta U = 0 = Q - W$, so $W = mgh = Q$. Hence you must climb to height $h = Q/mg$. First convert units: $Q = (900\text{ kcal})(4186\text{ J/1 kcal}) = 3.77 \times 10^6$ J. Then

$$h = \frac{Q}{mg} = \frac{3.77 \times 10^6\text{ J}}{(60.0\text{ kg})(9.80\text{ m/s}^2)} = 6410\text{ m}$$

EVALUATE We have unrealistically assumed 100% efficiency in the conversion of food energy into mechanical work. The actual efficiency is roughly 25%, so the work W you do as you "burn off" the sundae is only about $(0.25)(900\text{ kcal}) = 225$ kcal. (The remaining 75%, or 675 kcal, is transferred to your surroundings as heat.) Hence you actually must climb about $(0.25)(6410\text{ m}) = 1600$ m, or one *mile*! Do you really want that sundae?

KEYCONCEPT The first law of thermodynamics states that the change in internal energy ΔU of a system equals the heat flow Q into the system minus the work W that the system does on its surroundings. Q is positive if heat flows into the system, negative if heat flows out; W is positive if the system does work on its surroundings, negative if the surroundings do work on the system.

EXAMPLE 19.3 A cyclic process

WITH VARIATION PROBLEMS

Figure 19.12 shows a pV-diagram for a *cyclic* process in which the initial and final states of some thermodynamic system are the same. The state of the system starts at point a and proceeds counterclockwise in the pV-diagram to point b, then back to a; the total work is $W = -500$ J. (a) Why is the work negative? (b) Find the change in internal energy and the heat added during this process.

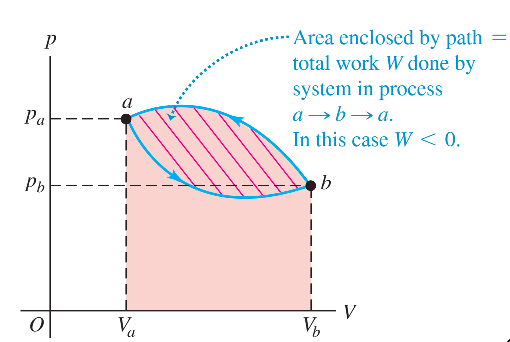


Figure 19.12 The net work done by the system in process $a \rightarrow b \rightarrow a$ is -500 J. What would it have been if the process had proceeded clockwise in this pV-diagram?

Continued

IDENTIFY and SET UP We must relate the change in internal energy, the heat added, and the work done in a thermodynamic process. Hence we can apply the first law of thermodynamics. The process is cyclic, and it has two steps: $a \rightarrow b$ via the lower curve in Fig. 19.12 and $b \rightarrow a$ via the upper curve. We are asked only about the *entire* cyclic process $a \rightarrow b \rightarrow a$.

EXECUTE (a) The work done in any step equals the area under the curve in the pV-diagram, with the area taken as positive if $V_2 > V_1$ and negative if $V_2 < V_1$; this rule yields the signs that result from the actual integrations in Eq. (19.2), $W = \int_{V_1}^{V_2} p\,dV$. The area under the lower curve $a \rightarrow b$ is therefore positive, but it is smaller than the absolute value of the (negative) area under the upper curve $b \rightarrow a$. Therefore the net area (the area enclosed by the path, shown with red stripes) and the net work W are negative. In other words, 500 J more work is done *on* the system than *by* the system in the complete process.

(b) In any cyclic process, $\Delta U = 0$, so $Q = W$. Here, that means $Q = -500$ J; that is, 500 J of heat flows *out of* the system.

EVALUATE In cyclic processes, the total work is positive if the process goes clockwise around the pV-diagram representing the cycle, and negative if the process goes counterclockwise (as here).

KEYCONCEPT In a cyclic thermodynamic process a system returns to the same state it was in initially, so the net change ΔU in internal energy is zero. The net work W done by the system in a cyclic process equals the area enclosed by the path that the system follows on a pV-diagram. Since $\Delta U = 0$, the net heat flow Q into the system in a cyclic process is equal to W.

EXAMPLE 19.4 Comparing thermodynamic processes

WITH VARIATION PROBLEMS

The pV-diagram of **Fig. 19.13** shows a series of thermodynamic processes. In process ab, 150 J of heat is added to the system; in process bd, 600 J of heat is added. Find (a) the internal energy change in process ab; (b) the internal energy change in process abd (shown in light blue); and (c) the total heat added in process acd (shown in dark blue).

IDENTIFY and SET UP In each process we use $\Delta U = Q - W$ to determine the desired quantity. We are given $Q_{ab} = +150$ J and $Q_{bd} = +600$ J (both values are positive because heat is *added* to the system). Our target variables are (a) ΔU_{ab}, (b) ΔU_{abd}, and (c) Q_{acd}.

EXECUTE (a) No volume change occurs during process ab, so the system does no work: $W_{ab} = 0$ and so $\Delta U_{ab} = Q_{ab} = 150$ J.

(b) Process bd is an expansion at constant pressure, so from Eq. (19.3),

$$\begin{aligned} W_{bd} &= p(V_2 - V_1) \\ &= (8.0 \times 10^4\ \text{Pa})(5.0 \times 10^{-3}\ \text{m}^3 - 2.0 \times 10^{-3}\ \text{m}^3) = 240\ \text{J} \end{aligned}$$

The total work for the two-step process abd is then

$$W_{abd} = W_{ab} + W_{bd} = 0 + 240\ \text{J} = 240\ \text{J}$$

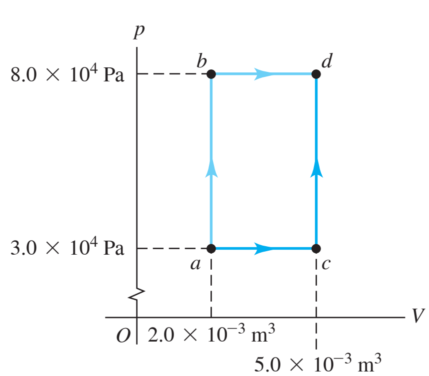
and the total heat is

$$Q_{abd} = Q_{ab} + Q_{bd} = 150\ \text{J} + 600\ \text{J} = 750\ \text{J}$$

Applying Eq. (19.4) to abd, we then have

$$\Delta U_{abd} = Q_{abd} - W_{abd} = 750\ \text{J} - 240\ \text{J} = 510\ \text{J}$$

Figure 19.13 A pV-diagram showing the various thermodynamic processes.



(c) Because ΔU is *independent of the path* from a to d, the internal energy change is the same for path acd as for path abd:

$$\Delta U_{acd} = \Delta U_{abd} = 510\ \text{J}$$

The total work for path acd is

$$\begin{aligned} W_{acd} &= W_{ac} + W_{cd} = p(V_2 - V_1) + 0 \\ &= (3.0 \times 10^4\ \text{Pa})(5.0 \times 10^{-3}\ \text{m}^3 - 2.0 \times 10^{-3}\ \text{m}^3) \\ &= 90\ \text{J} \end{aligned}$$

Now we apply Eq. (19.5) to process acd:

$$Q_{acd} = \Delta U_{acd} + W_{acd} = 510\ \text{J} + 90\ \text{J} = 600\ \text{J}$$

We tabulate the quantities above:

Step	Q	W	$\Delta U = Q - W$	Step	Q	W	$\Delta U = Q - W$
ab	150 J	0 J	150 J	ac	?	90 J	?
bd	600 J	240 J	360 J	cd	?	0 J	?
abd	750 J	240 J	510 J	acd	600 J	90 J	510 J

EVALUATE Be sure that you understand how each entry in the table above was determined. Although ΔU is the same (510 J) for abd and acd, W (240 J versus 90 J) and Q (750 J versus 600 J) are quite different. Although we couldn't find Q or ΔU for processes ac and cd, we could analyze the composite process acd by comparing it with process abd, which has the same initial and final states and for which we have more information.

KEYCONCEPT When a system starts in one state and ends in a different state, the internal energy change ΔU of the system is the same no matter what path the system takes between the two states. However, the heat flow Q into the system and the work W done by the system *do* depend on the path taken.

EXAMPLE 19.5 Thermodynamics of boiling water

WITH VARIATION PROBLEMS

One gram of water ($1\ \text{cm}^3$) becomes $1671\ \text{cm}^3$ of steam when boiled at a constant pressure of 1 atm (1.013×10^5 Pa). The heat of vaporization at this pressure is $L_v = 2.256 \times 10^6$ J/kg. Compute (a) the work done by the water when it vaporizes and (b) its increase in internal energy.

IDENTIFY and SET UP The heat added causes the system (water) to change phase from liquid to vapor. We can analyze this process by using the first law of thermodynamics. The water is boiled at constant pressure, so we can use Eq. (19.3) to calculate the work W done by the vaporizing water as it expands. We are given the mass of water and the heat of vaporization, so we can use Eq. (17.20), $Q = mL_v$, to calculate the heat Q added to the water. We can then find the internal energy change from Eq. (19.4), $\Delta U = Q - W$.

EXECUTE (a) From Eq. (19.3), the water does work

$$W = p(V_2 - V_1)$$
$$= (1.013 \times 10^5\ \text{Pa})(1671 \times 10^{-6}\ \text{m}^3 - 1 \times 10^{-6}\ \text{m}^3) = 169\ \text{J}$$

(b) From Eq. (17.20), the heat added to the water is

$$Q = mL_v = (10^{-3}\ \text{kg})(2.256 \times 10^6\ \text{J/kg}) = 2256\ \text{J}$$

Then from Eq. (19.4),

$$\Delta U = Q - W = 2256\ \text{J} - 169\ \text{J} = 2087\ \text{J}$$

EVALUATE To vaporize 1 g of water, we must add 2256 J of heat, most of which (2087 J) remains in the system as an increase in internal energy. The remaining 169 J leaves the system as the system expands from liquid to vapor and does work against the surroundings. (The increase in internal energy is associated mostly with the attractive intermolecular forces. The associated potential energies are greater after work has been done to pull apart the molecules in the liquid, forming the vapor state. It's like increasing gravitational potential energy by pulling an elevator farther from the center of the earth.)

KEYCONCEPT For a system that undergoes a thermodynamic process, if you can calculate any two of the quantities ΔU (internal energy change), Q (heat flow into the system), and W (work done by the system), you can calculate the third using the first law of thermodynamics, $\Delta U = Q - W$.

Infinitesimal Changes of State

In the preceding examples the initial and final states differ by a finite amount. Later we'll consider *infinitesimal* changes of state in which a small amount of heat dQ is added to the system, the system does a small amount of work dW, and its internal energy changes by an amount dU. For such a process,

First law of thermodynamics, infinitesimal process:

$$\underbrace{dU}_{\text{Infinitesimal internal energy change}} = \underbrace{dQ}_{\text{Infinitesimal heat added}} - \underbrace{dW}_{\text{Infinitesimal work done}} \qquad (19.6)$$

For the systems we'll discuss, the work dW is given by $dW = p\,dV$, so we can also state the first law as

$$dU = dQ - p\,dV \qquad (19.7)$$

TEST YOUR UNDERSTANDING OF SECTION 19.4 Rank the following thermodynamic processes according to the change in internal energy in each process, from most positive to most negative. (i) As you do 250 J of work on a system, it transfers 250 J of heat to its surroundings; (ii) as you do 250 J of work on a system, it absorbs 250 J of heat from its surroundings; (iii) as a system does 250 J of work on you, it transfers 250 J of heat to its surroundings; (iv) as a system does 250 J of work on you, it absorbs 250 J of heat from its surroundings.

ANSWER

(ii), (i) and (iv) (tie), (iii) In the expression $\Delta U = Q - W$, Q is the heat *added* to the system and W is the work done *by* the system. If heat is transferred from the system to its surroundings, Q is negative; if work is done on the system, W is negative. Hence we have (i) $Q = -250$ J, $W = -250$ J, $\Delta U = -250\ \text{J} - (-250\ \text{J}) = 0$; (ii) $Q = 250$ J, $W = -250$ J, $\Delta U = 250\ \text{J} - (-250\ \text{J}) = 500$ J; (iii) $Q = -250$ J, $W = 250$ J, $\Delta U = -250\ \text{J} - 250\ \text{J} = -500$ J; and (iv) $Q = 250$ J, $W = 250$ J, $\Delta U = 250\ \text{J} - 250\ \text{J} = 0$.

19.5 KINDS OF THERMODYNAMIC PROCESSES

In this section we describe four specific kinds of thermodynamic processes that occur often in practical situations. We can summarize these briefly as "no heat transfer" or *adiabatic,* "constant volume" or *isochoric,* "constant pressure" or *isobaric,* and "constant temperature" or *isothermal*. For some of these processes we can use a simplified form of the first law of thermodynamics.

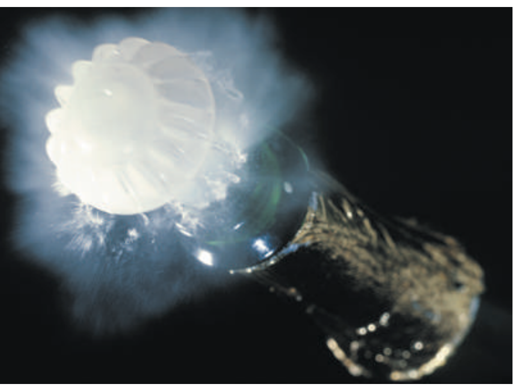
Adiabatic Process

An **adiabatic process** (pronounced "ay-dee-ah-*bat*-ic") is defined as one with no heat transfer into or out of a system; $Q = 0$. We can prevent heat flow either by surrounding the system with thermally insulating material or by carrying out the process so quickly that there is not enough time for appreciable heat flow. From the first law we find that for every adiabatic process,

$$U_2 - U_1 = \Delta U = -W \qquad \text{(adiabatic process)} \tag{19.8}$$

When a system expands adiabatically, W is positive (the system does work on its surroundings), so ΔU is negative and the internal energy decreases. When a system is *compressed* adiabatically, W is negative (work is done on the system by its surroundings) and U increases. In many (but not all) systems an increase of internal energy is accompanied by a rise in temperature, and a decrease in internal energy by a drop in temperature (**Fig. 19.14**).

Figure 19.14 When the cork is popped on a bottle of champagne, the pressurized gases inside the bottle expand rapidly and do work on the outside air ($W > 0$). There is little time for the gases to exchange heat with their surroundings, so the expansion is nearly adiabatic ($Q = 0$). Hence the internal energy of the expanding gases decreases ($\Delta U = -W < 0$) and their temperature drops. This makes water vapor condense and form a miniature cloud.



The compression stroke in an internal-combustion engine is an approximately adiabatic process. The temperature rises as the air–fuel mixture in the cylinder is compressed. The expansion of the burned fuel during the power stroke is also an approximately adiabatic expansion with a drop in temperature. In Section 19.8 we'll consider adiabatic processes in an ideal gas.

Isochoric Process

An **isochoric process** (pronounced "eye-so-*kor*-ic") is a *constant-volume* process. When the volume of a thermodynamic system is constant, it does no work on its surroundings. Then $W = 0$ and

$$U_2 - U_1 = \Delta U = Q \qquad \text{(isochoric process)} \tag{19.9}$$

In an isochoric process, all the energy added as heat remains in the system as an increase in internal energy. Heating a gas in a closed constant-volume container is an example of an isochoric process. The processes *ab* and *cd* in Example 19.4 are also examples of isochoric processes. (Note that there are types of work that do not involve a volume change. For example, we can do work on a fluid by stirring it. In some literature, "isochoric" is used to mean that no work of any kind is done.)

Figure 19.15 Most cooking involves isobaric processes. That's because the air pressure above a saucepan or frying pan, or inside a microwave oven, remains essentially constant while the food is being heated.

Isobaric Process

An **isobaric process** (pronounced "eye-so-*bear*-ic") is a *constant-pressure* process. In general, none of the three quantities ΔU, Q, and W is zero in an isobaric process, but calculating W is easy nonetheless. From Eq. (19.3),

$$W = p(V_2 - V_1) \qquad \text{(isobaric process)} \tag{19.10}$$

Boiling water at constant pressure is an isobaric process (**Fig. 19.15**).

Isothermal Process

An **isothermal process** is a *constant-temperature* process. For a process to be isothermal, any heat flow into or out of the system must occur slowly enough that thermal equilibrium is maintained. In general, none of the quantities ΔU, Q, or W is zero in an isothermal process.

In some special cases the internal energy of a system depends *only* on its temperature, not on its pressure or volume. The most familiar system having this special property is an ideal gas, as we'll discuss in the next section. For such systems, if the temperature is constant, the internal energy is also constant; $\Delta U = 0$ and $Q = W$. That is, any energy entering the system as heat Q must leave it again as work W done by the system. Example 19.1, involving an ideal gas, is an example of an isothermal process in which U is also constant. For most systems other than ideal gases, the internal energy depends on pressure as well as temperature, so U may vary even when T is constant.

Figure 19.16 shows a pV-diagram for these four processes for a constant amount of an ideal gas. The path followed in an adiabatic process (a to 1) is called an **adiabat.** A vertical line (constant volume) is an **isochor,** a horizontal line (constant pressure) is an **isobar,** and a curve of constant temperature (shown as light blue lines in Fig. 19.16) is an **isotherm.**

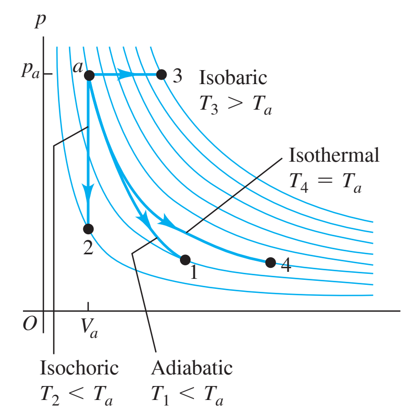


Figure 19.16 Four different processes for a constant amount of an ideal gas, all starting at state a. For the adiabatic process, $Q = 0$; for the isochoric process, $W = 0$; and for the isothermal process, $\Delta U = 0$. The temperature increases only during the isobaric expansion.

TEST YOUR UNDERSTANDING OF SECTION 19.5 Which of the processes in Fig. 19.7 are isochoric? Which are isobaric? Is it possible to tell if any of the processes are isothermal or adiabatic?

ANSWER

❙ **1 → 4 and 3 → 2 are isochoric; 1 → 3 and 4 → 2 are isobaric; no** In a pV-diagram like those shown in Fig. 19.7, isochoric processes are represented by vertical lines (lines of constant volume) and isobaric processes are represented by horizontal lines (lines of constant pressure). The process 1 → 2 in Fig. 19.7 is shown as a curved line, which superficially resembles the adiabatic and isothermal processes for an ideal gas in Fig. 19.16. Without more information we can't tell whether process 1 → 2 is isothermal, adiabatic, or neither.

19.6 INTERNAL ENERGY OF AN IDEAL GAS

We now show that for an ideal gas, the internal energy U depends only on temperature, not on pressure or volume. Let's think again about the free-expansion experiment described in Section 19.3. A thermally insulated container with rigid walls is divided into two compartments by a partition (**Fig. 19.17**). One compartment has a quantity of an ideal gas and the other is evacuated.

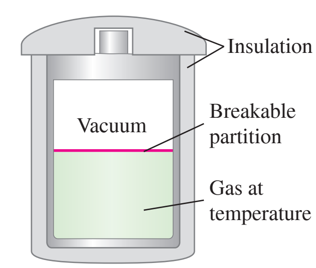


Figure 19.17 The partition is broken (or removed) to start the free expansion of gas into the vacuum region.

When the partition is removed or broken, the gas expands to fill both parts of the container. There is no heat flow through the insulation, and the gas does no work on its surroundings because the walls of the container don't move. So both Q and W are zero and the internal energy U is constant.

Does the *temperature* T change during a free expansion? Suppose it *does* change, while the internal energy stays the same. In that case we have to conclude that the internal energy depends on both T and the volume V or on both T and the pressure p, but certainly not on T alone. But if T is constant during a free expansion, for which we know that U is constant even though both p and V change, then we have to conclude that U depends only on T, not on p or V.

Many experiments have shown that when a low-density gas (essentially an ideal gas) undergoes a free expansion, its temperature *does not* change. The conclusion is:

The internal energy U of an ideal gas depends only on its temperature T, not on its pressure or volume.

This property, in addition to the ideal-gas equation of state, is part of the ideal-gas model. We'll make frequent use of this property.

For nonideal gases, some temperature change occurs during free expansions, even though the internal energy is constant. This shows that the internal energy cannot depend *only* on temperature; it must depend on pressure as well. From the microscopic viewpoint, in which internal energy U is the sum of the kinetic and potential energies for all the particles that make up the system, this is not surprising. Nonideal gases usually have attractive intermolecular forces, and when molecules move farther apart, the associated potential energies increase. If the total internal energy is constant, the kinetic energies must decrease. Temperature is directly related to molecular *kinetic* energy, and for such a gas a free expansion is usually accompanied by a *drop* in temperature.

TEST YOUR UNDERSTANDING OF SECTION 19.6 Is the internal energy of a solid likely to be independent of its volume, as is the case for an ideal gas? Explain your reasoning. (*Hint:* See Fig. 18.20.)

ANSWER

❙ **no** Using the model of a solid in Fig. 18.20, we can see that the internal energy of a solid *does* depend on its volume. Compressing the solid means compressing the "springs" between the atoms, thereby increasing their stored potential energy and hence the internal energy of the solid.

Figure **19.18** Measuring the molar heat capacity of an ideal gas (a) at constant volume and (b) at constant pressure.

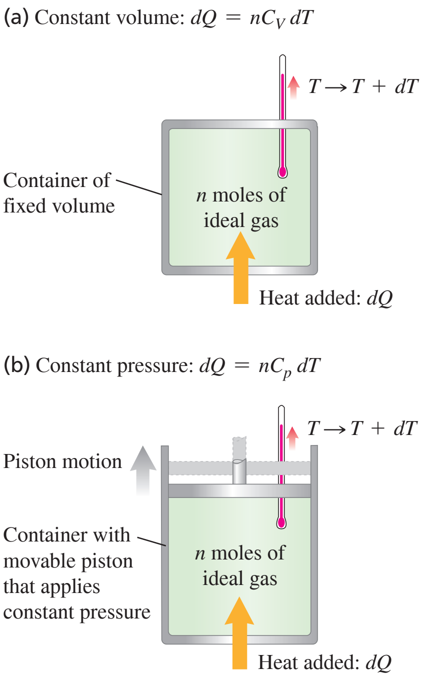
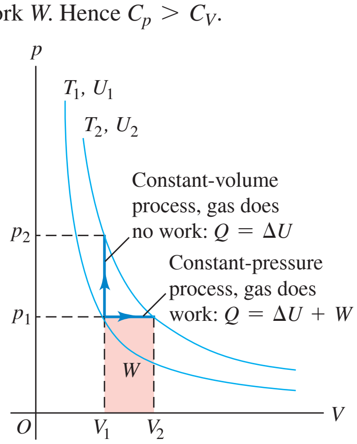


Figure **19.19** Raising the temperature of an ideal gas from T_1 to T_2 by a constant-volume or a constant-pressure process. For an ideal gas, U depends only on T, so ΔU is the same for both processes. But for the constant-pressure process, more heat Q must be added to both increase U and do work W. Hence $C_p > C_V$.



19.7 HEAT CAPACITIES OF AN IDEAL GAS

We defined specific heat and molar heat capacity in Section 17.5. We also remarked at the end of that section that the specific heat or molar heat capacity of a substance depends on the conditions under which the heat is added. The heat capacity of a gas is usually measured in a closed container under constant-volume conditions. The corresponding heat capacity is the **molar heat capacity at constant volume,** denoted by C_V. Heat capacity measurements for solids and liquids are usually carried out under constant atmospheric pressure, and we call the corresponding heat capacity the **molar heat capacity at constant pressure,** C_p.

Let's consider C_V and C_p for an ideal gas. To measure C_V, we raise the temperature of an ideal gas in a rigid container with constant volume, ignoring its thermal expansion (**Fig. 19.18a**). To measure C_p, we let the gas expand just enough to keep the pressure constant as the temperature rises (Fig. 19.18b).

Why should these two molar heat capacities be different? The answer lies in the first law of thermodynamics. In a constant-volume temperature increase, the system does no work, and the change in internal energy ΔU equals the heat added Q. In a constant-pressure temperature increase, on the other hand, the volume *must* increase; otherwise, the pressure (given by the ideal-gas equation of state, $p = nRT/V$) could not remain constant. As the material expands, it does an amount of work W. According to the first law,

$$Q = \Delta U + W \qquad (19.11)$$

For a given temperature increase, the internal energy change ΔU of an ideal gas has the same value no matter what the process (remember that the internal energy of an ideal gas depends only on temperature, not on pressure or volume). Equation (19.11) then shows that the heat input for a constant-pressure process must be *greater* than that for a constant-volume process because additional energy must be supplied to account for the work W done during the expansion. So C_p is greater than C_V for an ideal gas. The pV-diagram in **Fig. 19.19** shows this relationship. For air, C_p is 40% greater than C_V.

For a very few substances (one of which is water between 0°C and 4°C) the volume *decreases* during heating. In this case, W is negative and the internal energy change ΔU is greater than the heat input Q.

Relating C_p and C_V for an Ideal Gas

We can derive a simple relationship between C_p and C_V for an ideal gas. First consider the constant-*volume* process. We place n moles of an ideal gas at temperature T in a constant-volume container. We place it in thermal contact with a hotter object; an infinitesimal quantity of heat dQ flows into the gas, and its temperature increases by an infinitesimal amount dT. By the definition of C_V, the molar heat capacity at constant volume,

$$dQ = nC_V\,dT \qquad (19.12)$$

The pressure increases during this process, but the gas does no work $(dW = 0)$ because the volume is constant. The first law in differential form, Eq. (19.6), is $dQ = dU + dW$. Since $dW = 0$, $dQ = dU$ and Eq. (19.12) can also be written as

$$dU = nC_V\,dT \qquad (19.13)$$

Now consider a constant-*pressure* process with the same temperature change dT. We place the same gas in a cylinder with a piston that we allow to move just enough to maintain constant pressure (Fig. 19.18b). Again we bring the system into contact with a hotter object. As heat flows into the gas, it expands at constant pressure and does work. By the definition of C_p, the molar heat capacity at constant pressure, the amount of heat entering the gas is

$$dQ = nC_p\,dT \qquad (19.14)$$

The work dW done by the gas in this constant-pressure process is

$$dW = p\,dV$$

We can also express dW in terms of the temperature change dT by using the ideal-gas equation of state, $pV = nRT$. Because p is constant, the change in V is proportional to the change in T:

$$dW = p\,dV = nR\,dT \tag{19.15}$$

Now substitute Eqs. (19.14) and (19.15) into the first law, $dQ = dU + dW$:

$$nC_p\,dT = dU + nR\,dT \tag{19.16}$$

Here comes the crux of the calculation. The internal energy change dU for the constant-pressure process is again given by Eq. (19.13), $dU = nC_V\,dT$, *even though now the volume is not constant.* Why is this so? From Section 19.6, one of the special properties of an ideal gas is that its internal energy depends *only* on temperature. Thus the *change* in internal energy during any process must be determined only by the temperature change. If Eq. (19.13) is valid for an ideal gas for one particular kind of process, it must be valid for an ideal gas for *every* kind of process with the same dT. So we may replace dU in Eq. (19.16) by $nC_V\,dT$:

$$nC_p\,dT = nC_V\,dT + nR\,dT$$

When we divide each term by the common factor $n\,dT$, we get

For an ideal gas:

$$\underbrace{C_p}_{\text{Molar heat capacity at constant } pressure} = \underbrace{C_V}_{\text{Molar heat capacity at constant } volume} + \underbrace{R}_{\text{Gas constant}} \tag{19.17}$$

As we predicted, the molar heat capacity of an ideal gas at constant pressure is *greater* than the molar heat capacity at constant volume; the difference is the gas constant R.

We have used the ideal-gas model to derive Eq. (19.17), but it turns out to be obeyed to within a few percent by many real gases at moderate pressures. **Table 19.1** gives measured values of C_p and C_V for several real gases at low pressures; the difference in most cases is approximately $R = 8.314\ \text{J/mol}\cdot\text{K}$.

The table also shows that the molar heat capacity of a gas is related to its molecular structure, as we discussed in Section 18.4. In fact, the first three columns of Table 19.1 are the same as Table 18.1.

The Ratio of Heat Capacities

The last column of Table 19.1 lists the values of the dimensionless **ratio of heat capacities, C_p/C_V**, denoted by γ (the Greek letter gamma):

$$\underbrace{\gamma}_{\text{Ratio of heat capacities}} = \frac{\overbrace{C_p}^{\text{Molar heat capacity at constant } pressure}}{\underbrace{C_V}_{\text{Molar heat capacity at constant } volume}} \tag{19.18}$$

TABLE 19.1 Molar Heat Capacities of Gases at Low Pressure

Type of Gas	Gas	C_V (J/mol · K)	C_p (J/mol · K)	$C_p - C_V$ (J/mol · K)	$\gamma = C_p/C_V$
Monatomic	He	12.47	20.78	8.31	1.67
	Ar	12.47	20.78	8.31	1.67
Diatomic	H_2	20.42	28.74	8.32	1.41
	N_2	20.76	29.07	8.31	1.40
	O_2	20.85	29.17	8.32	1.40
	CO	20.85	29.16	8.31	1.40
Polyatomic	CO_2	28.46	36.94	8.48	1.30
	SO_2	31.39	40.37	8.98	1.29
	H_2S	25.95	34.60	8.65	1.33

(This is sometimes called the "ratio of specific heats.") For gases, C_p is always greater than C_V and γ is always greater than unity. We'll see in the next section that γ plays an important role in adiabatic processes for an ideal gas.

We can use our kinetic-theory discussion of the molar heat capacity of an ideal gas (see Section 18.4) to predict values of γ. As an example, an ideal monatomic gas has $C_V = \frac{3}{2}R$. From Eq. (19.17),

$$C_p = C_V + R = \tfrac{3}{2}R + R = \tfrac{5}{2}R$$

so

$$\gamma = \frac{C_p}{C_V} = \frac{\frac{5}{2}R}{\frac{3}{2}R} = \frac{5}{3} = 1.67$$

As Table 19.1 shows, this agrees well with values of γ computed from measured heat capacities. For most diatomic gases near room temperature, $C_V = \frac{5}{2}R$, $C_p = C_V + R = \frac{7}{2}R$, and

$$\gamma = \frac{C_p}{C_V} = \frac{\frac{7}{2}R}{\frac{5}{2}R} = \frac{7}{5} = 1.40$$

also in good agreement with measured values.

Here's a final reminder: For an ideal gas, the internal energy change in *any* process is given by $\Delta U = nC_V\,\Delta T$, *whether the volume is constant or not.* This relationship holds for other substances *only* when the volume is constant.

CAUTION **A few hints about ΔU, Q, and W for an ideal gas** Remember that if the gas temperature T increases, the internal energy change ΔU of the gas is positive; if T decreases, ΔU is negative. Remember, too, that the work W done by the gas is positive if the gas expands and negative if the gas is compressed. Finally, you can calculate the heat flow Q into the gas directly only in certain special processes (isochoric, isobaric, or adiabatic processes); for other processes, you must first calculate ΔU and W and then find Q using the first law of thermodynamics.

EXAMPLE 19.6 Cooling your room

WITH VARIATION PROBLEMS

A typical dorm room or bedroom contains about 2500 moles of air. Find the change in the internal energy of this much air when it is cooled from 35.0°C to 26.0°C at a constant pressure of 1.00 atm. Treat the air as an ideal gas with $\gamma = 1.400$.

IDENTIFY and SET UP Our target variable is the change in the internal energy ΔU of an ideal gas in a constant-pressure process. We are given the number of moles, the temperature change, and the value of γ for air. We use Eq. (19.13), $\Delta U = nC_V\,\Delta T$, which gives the internal energy change for an ideal gas in *any* process, *whether the volume is constant or not*. [See the discussion following Eq. (19.16).] We use Eqs. (19.17) and (19.18) to find C_V.

EXECUTE From Eqs. (19.17) and (19.18),

$$\gamma = \frac{C_p}{C_V} = \frac{C_V + R}{C_V} = 1 + \frac{R}{C_V}$$

$$C_V = \frac{R}{\gamma - 1} = \frac{8.314\ \text{J/mol}\cdot\text{K}}{1.400 - 1} = 20.79\ \text{J/mol}\cdot\text{K}$$

Then from Eq. (19.13),

$$\begin{aligned}\Delta U &= nC_V\,\Delta T\\ &= (2500\ \text{mol})(20.79\ \text{J/mol}\cdot\text{K})(26.0°\text{C} - 35.0°\text{C})\\ &= -4.68 \times 10^5\ \text{J}\end{aligned}$$

EVALUATE To cool 2500 moles of air from 35.0°C to 26.0°C, a room air conditioner must extract this much internal energy from the air and transfer it to the air outside. In Chapter 20 we'll discuss how this is done.

KEYCONCEPT When n moles of an ideal gas undergo a thermodynamic process in which the temperature changes by ΔT, the change in internal energy is equal to the product of n, ΔT, and the molar heat capacity at constant volume C_V: $\Delta U = nC_V\Delta T$. This is true whether or not the volume of the system remains constant.

TEST YOUR UNDERSTANDING OF SECTION 19.7 You want to cool a storage cylinder containing 10 moles of compressed gas from 30°C to 20°C. For which kind of gas would this be easiest? (i) A monatomic gas; (ii) a diatomic gas; (iii) a polyatomic gas; (iv) it would be equally easy for all of these.

ANSWER

(i) For a given number of moles n and a given temperature change ΔT, the amount of heat that must be transferred out of a fixed volume of air is $Q = nC_V\Delta T$. Hence the amount of heat transfer required is least for the gas with the smallest value of C_V. From Table 19.1, C_V is smallest for monatomic gases.

19.8 ADIABATIC PROCESSES FOR AN IDEAL GAS

An adiabatic process, defined in Section 19.5, is a process in which no heat transfer takes place between a system and its surroundings. Zero heat transfer is an idealization, but a process is approximately adiabatic if the system is well insulated or if the process takes place so quickly that there is not enough time for appreciable heat flow to occur.

In an adiabatic process, $Q = 0$, so from the first law, $\Delta U = -W$. An adiabatic process for an ideal gas is shown in the pV-diagram of **Fig. 19.20**. As the gas expands from volume V_a to V_b, it does positive work, so its internal energy decreases and its temperature drops. If point a, representing the initial state, lies on an isotherm at temperature $T + dT$, then point b for the final state is on a different isotherm at a lower temperature T. For an adiabatic *compression* from V_b to V_a the situation is reversed and the temperature rises.

The air in the output hoses of air compressors used to inflate tires and to fill scuba tanks is always warmer than the air entering the compressor; this is because the compression is rapid and hence approximately adiabatic. Adiabatic *cooling* occurs when you open a bottle of your favorite carbonated beverage. The gas just above the beverage surface expands rapidly in a nearly adiabatic process; the gas temperature drops so much that water vapor in the gas condenses, forming a miniature cloud (see Fig. 19.14).

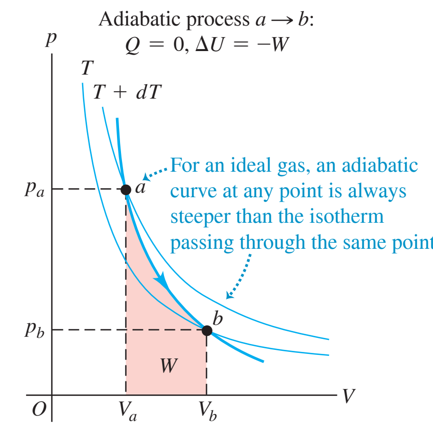


Figure 19.20 A pV-diagram of an adiabatic ($Q = 0$) process for an ideal gas. As the gas expands from V_a to V_b, it does positive work W on its environment, its internal energy decreases ($\Delta U = -W < 0$), and its temperature drops from $T + dT$ to T. (An adiabatic process is also shown in Fig. 19.16.)

CAUTION **"Heating" and "cooling" without heat** When we talk about "adiabatic heating" and "adiabatic cooling," we really mean "raising the temperature" and "lowering the temperature," respectively. In an adiabatic process, the temperature change is due to work done by or on the system; there is *no* heat flow at all.

Adiabatic Ideal Gas: Relating *V*, *T*, and *p*

We can derive a relationship between volume and temperature changes for an infinitesimal adiabatic process in an ideal gas. Equation (19.13) gives the internal energy change dU for *any* process for an ideal gas, adiabatic or not, so we have $dU = nC_V\,dT$. Also, the work done by the gas during the process is given by $dW = p\,dV$. Then, since $dU = -dW$ for an adiabatic process, we have

$$nC_V\,dT = -p\,dV \tag{19.19}$$

To obtain a relationship containing only the volume V and temperature T, we eliminate p by using the ideal-gas equation in the form $p = nRT/V$. Substituting this into Eq. (19.19) and rearranging, we get

$$nC_V\,dT = -\frac{nRT}{V}dV$$

$$\frac{dT}{T} + \frac{R}{C_V}\frac{dV}{V} = 0$$

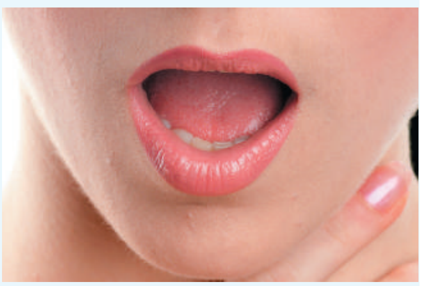
The coefficient R/C_V can be expressed in terms of $\gamma = C_p/C_V$. We have

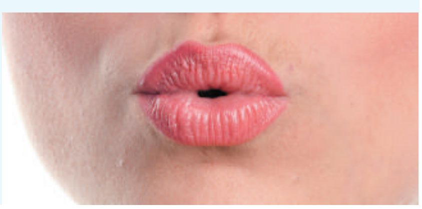
$$\frac{R}{C_V} = \frac{C_p - C_V}{C_V} = \frac{C_p}{C_V} - 1 = \gamma - 1$$

$$\frac{dT}{T} + (\gamma - 1)\frac{dV}{V} = 0 \tag{19.20}$$

Because γ is always greater than unity for a gas, $(\gamma - 1)$ is always positive. This means that in Eq. (19.20), dV and dT always have opposite signs. An adiabatic *expansion* of an ideal gas ($dV > 0$) always occurs with a *drop* in temperature ($dT < 0$), and an adiabatic *compression* ($dV < 0$) always occurs with a *rise* in temperature ($dT > 0$); this confirms our earlier prediction.

BIO APPLICATION Exhaling Adiabatically Put your hand a few centimeters in front of your mouth, open your mouth wide, and exhale. Your breath will feel warm on your hand, because the exhaled gases emerge at roughly the temperature of your body's interior. Now purse your lips as though you were going to whistle, and again blow on your hand. The exhaled gases will feel much cooler. In this case the gases undergo a rapid, essentially adiabatic expansion as they emerge from between your lips, so the temperature of the exhaled gases decreases.





For finite changes in temperature and volume we integrate Eq. (19.20), obtaining

$$\ln T + (\gamma - 1)\ln V = \text{constant}$$
$$\ln T + \ln V^{\gamma-1} = \text{constant}$$
$$\ln\left(TV^{\gamma-1}\right) = \text{constant}$$

and finally,

$$TV^{\gamma-1} = \text{constant} \tag{19.21}$$

Thus for an initial state (T_1, V_1) and a final state (T_2, V_2),

$$T_1V_1^{\gamma-1} = T_2V_2^{\gamma-1} \qquad \text{(adiabatic process, ideal gas)} \tag{19.22}$$

Because we have used the ideal-gas equation in our derivation of Eqs. (19.21) and (19.22), the T's must always be *absolute* (Kelvin) temperatures.

We can also convert Eq. (19.21) into a relationship between pressure and volume by eliminating T, using the ideal-gas equation in the form $T = pV/nR$. Substituting this into Eq. (19.21), we find

$$\frac{pV}{nR}V^{\gamma-1} = \text{constant}$$

or, because n and R are constant,

$$pV^{\gamma} = \text{constant} \tag{19.23}$$

For an initial state (p_1, V_1) and a final state (p_2, V_2), Eq. (19.23) becomes

$$p_1V_1^{\gamma} = p_2V_2^{\gamma} \qquad \text{(adiabatic process, ideal gas)} \tag{19.24}$$

We can also calculate the *work* done by an ideal gas during an adiabatic process. We know that $Q = 0$ and $W = -\Delta U$ for *any* adiabatic process. For an ideal gas, $\Delta U = nC_V(T_2 - T_1)$. If the number of moles n and the initial and final temperatures T_1 and T_2 are known, we have simply

Work done by an ideal gas, adiabatic process

Number of moles; Initial temperature; Final temperature; Molar heat capacity at constant volume

$$W = nC_V(T_1 - T_2) \tag{19.25}$$

We may also use $pV = nRT$ in this equation to obtain

Work done by an ideal gas, adiabatic process

Molar heat capacity at constant volume; Gas constant; Ratio of heat capacities; Initial pressure, volume; Final pressure, volume

$$W = \frac{C_V}{R}(p_1V_1 - p_2V_2) = \frac{1}{\gamma - 1}(p_1V_1 - p_2V_2) \tag{19.26}$$

(We used the result $C_V = R/(\gamma - 1)$ from Example 19.6.) If the process is an expansion, the temperature drops, T_1 is greater than T_2, p_1V_1 is greater than p_2V_2, and the work is *positive*. If the process is a compression, the work is negative.

Throughout this analysis of adiabatic processes we have used the ideal-gas equation of state, which is valid only for *equilibrium* states. Strictly speaking, our results are valid only for a process that is fast enough to prevent appreciable heat exchange with the surroundings (so that $Q = 0$ and the process is adiabatic), yet slow enough that the system does not depart very much from thermal and mechanical equilibrium. Even when these conditions are not strictly satisfied, though, Eqs. (19.22), (19.24), and (19.26) give useful approximate results.

EXAMPLE 19.7 Adiabatic compression in a diesel engine

WITH VARIATION PROBLEMS

The compression ratio of a diesel engine is 15.0 to 1; that is, air in a cylinder is compressed to $\frac{1}{(15.0)}$ of its initial volume (**Fig. 19.21**). (a) If the initial pressure is 1.01×10^5 Pa and the initial temperature is 27°C (300 K), find the final pressure and the temperature after adiabatic compression. (b) How much work does the gas do during the compression if the initial volume of the cylinder is $1.00 \text{ L} = 1.00 \times 10^{-3} \text{ m}^3$? Use the values $C_V = 20.8 \text{ J/mol} \cdot \text{K}$ and $\gamma = 1.400$ for air.

IDENTIFY and SET UP This problem involves the adiabatic compression of an ideal gas, so we can use the ideas of this section. In part (a) we are given the initial pressure and temperature $p_1 = 1.01 \times 10^5$ Pa and $T_1 = 300$ K; the ratio of initial and final volumes is $V_1/V_2 = 15.0$. We use Eq. (19.22) to find the final temperature T_2 and Eq. (19.24) to find the final pressure p_2. In part (b) our target variable is W, the work done *by* the gas during the adiabatic compression. We use Eq. (19.26) to calculate W.

EXECUTE (a) From Eqs. (19.22) and (19.24),

$$T_2 = T_1\left(\frac{V_1}{V_2}\right)^{\gamma-1} = (300 \text{ K})(15.0)^{0.40} = 886 \text{ K} = 613°\text{C}$$

$$p_2 = p_1\left(\frac{V_1}{V_2}\right)^{\gamma} = (1.01 \times 10^5 \text{ Pa})(15.0)^{1.40}$$
$$= 44.8 \times 10^5 \text{ Pa} = 44 \text{ atm}$$

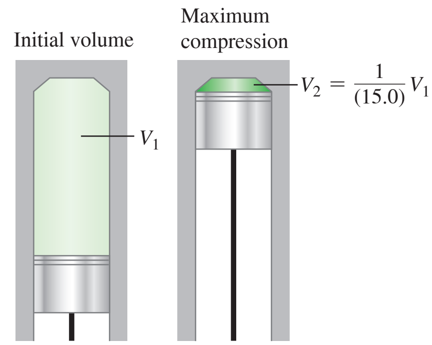
(b) From Eq. (19.26), the work done is

$$W = \frac{1}{\gamma - 1}(p_1V_1 - p_2V_2)$$

Using $V_1/V_2 = 15.0$, we have

$$W = \frac{1}{1.400 - 1}\left[\begin{array}{l}(1.01 \times 10^5 \text{ Pa})(1.00 \times 10^{-3} \text{ m}^3) \\ - (44.8 \times 10^5 \text{ Pa})\left(\dfrac{1.00 \times 10^{-3} \text{ m}^3}{15.0}\right)\end{array}\right]$$
$$= -494 \text{ J}$$

Figure **19.21** Adiabatic compression of air in a cylinder of a diesel engine.



EVALUATE If the compression had been isothermal, the final pressure would have been 15.0 atm. Because the temperature also increases during an adiabatic compression, the final pressure is much greater. When fuel is injected into the cylinders near the end of the compression stroke, the high temperature of the air attained during compression causes the fuel to ignite spontaneously without the need for spark plugs.

We can check our result in part (b) by using Eq. (19.25). The number of moles of gas in the cylinder is

$$n = \frac{p_1V_1}{RT_1} = \frac{(1.01 \times 10^5 \text{ Pa})(1.00 \times 10^{-3} \text{ m}^3)}{(8.314 \text{ J/mol} \cdot \text{K})(300 \text{ K})} = 0.0405 \text{ mol}$$

Then Eq. (19.25) gives

$$W = nC_V(T_1 - T_2)$$
$$= (0.0405 \text{ mol})(20.8 \text{ J/mol} \cdot \text{K})(300 \text{ K} - 886 \text{ K}) = -494 \text{ J}$$

The work is negative because the gas is compressed.

KEYCONCEPT In an adiabatic process there is no heat flow into or out of the system, so $Q = 0$ and the internal energy equals the negative of the work done by the system: $\Delta U = -W$. If the system is an ideal gas, the temperature T, volume V, and pressure p all change in the process, but the quantities $TV^{\gamma-1}$ and PV^{γ} remain constant (γ is the ratio of specific heats for the gas).

TEST YOUR UNDERSTANDING OF SECTION 19.8 You have four samples of ideal gas, each of which contains the same number of moles of gas and has the same initial temperature, volume, and pressure. You compress each sample to one-half of its initial volume. Rank the four samples in order from highest to lowest value of the final pressure. (i) A monatomic gas compressed isothermally; (ii) a monatomic gas compressed adiabatically; (iii) a diatomic gas compressed isothermally; (iv) a diatomic gas compressed adiabatically.

ANSWER

(ii), (iv), (i) and (iii) (tie) Samples (i) and (iii) are compressed isothermally, so pV = constant. The volume of each sample decreases to one-half of its initial value, so the final pressure is twice the initial pressure. Samples (ii) and (iv) are compressed adiabatically, so pV^{γ} = constant and the pressure increases by a factor of 2^{γ}. Sample (ii) is a monatomic gas for which $\gamma = \frac{5}{3}$, so its final pressure is $2^{5/3} = 3.17$ times greater than the initial pressure. Sample (iv) is a diatomic gas for which $\gamma = \frac{7}{5}$, so its final pressure is greater than the initial pressure by a factor of $2^{7/5} = 2.64$.

CHAPTER 19 SUMMARY

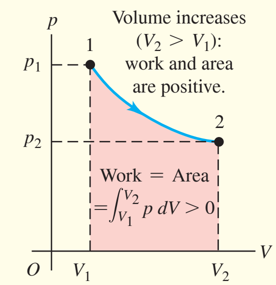
Heat and work in thermodynamic processes: A thermodynamic system has the potential to exchange energy with its surroundings by heat transfer or by mechanical work. When a system at pressure p changes volume from V_1 to V_2, it does an amount of work W given by the integral of p with respect to volume. If the pressure is constant, the work done is equal to p times the change in volume. A negative value of W means that work is done on the system. (See Example 19.1.)

In any thermodynamic process, the heat added to the system and the work done by the system depend not only on the initial and final states, but also on the path (the series of intermediate states through which the system passes).

$$W = \int_{V_1}^{V_2} p\, dV \qquad (19.2)$$

$$W = p(V_2 - V_1) \qquad (19.3)$$

(constant pressure only)



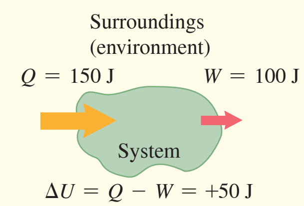
The first law of thermodynamics: The first law of thermodynamics states that when heat Q is added to a system while the system does work W, the internal energy U changes by an amount equal to $Q - W$. This law can also be expressed for an infinitesimal process. (See Examples 19.2, 19.3, and 19.5.)

The internal energy of any thermodynamic system depends only on its state. The change in internal energy in any process depends only on the initial and final states, not on the path. The internal energy of an isolated system is constant. (See Example 19.4.)

$$\Delta U = Q - W \qquad (19.4)$$

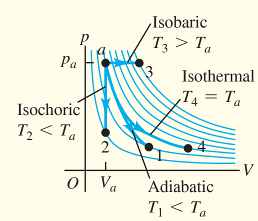
$$dU = dQ - dW \qquad (19.6)$$

(infinitesimal process)



Important kinds of thermodynamic processes:

- Adiabatic process: No heat transfer into or out of a system; $Q = 0$.
- Isochoric process: Constant volume; $W = 0$.
- Isobaric process: Constant pressure; $W = p(V_2 - V_1)$.
- Isothermal process: Constant temperature.

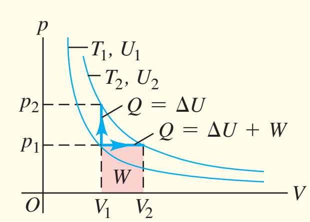


Thermodynamics of ideal gases: The internal energy of an ideal gas depends only on its temperature, not on its pressure or volume. For other substances the internal energy generally depends on both pressure and temperature.

The molar heat capacities C_V and C_p of an ideal gas differ by R, the ideal-gas constant. The dimensionless ratio of heat capacities, C_p/C_V, is denoted by γ. (See Example 19.6.)

$$C_p = C_V + R \qquad (19.17)$$

$$\gamma = \frac{C_p}{C_V} \qquad (19.18)$$

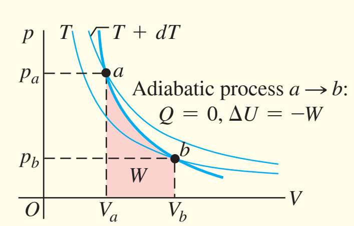


Adiabatic processes in ideal gases: For an adiabatic process for an ideal gas, the quantities $TV^{\gamma-1}$ and pV^{γ} are constant. The work done by an ideal gas during an adiabatic expansion can be expressed in terms of the initial and final values of temperature, or in terms of the initial and final values of pressure and volume. (See Example 19.7.)

$$W = nC_V(T_1 - T_2) \qquad (19.25)$$

$$= \frac{C_V}{R}(p_1V_1 - p_2V_2) \qquad (19.26)$$

$$= \frac{1}{\gamma - 1}(p_1V_1 - p_2V_2)$$



Chapter 19 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 19.2, 19.3, 19.4, and 19.5 (Section 19.4) before attempting these problems.

VP19.5.1 Consider the following thermodynamic processes. (a) The internal energy of a quantity of gas increases by 2.50×10^3 J as the gas does 3.25×10^3 J of work on its surroundings. How much heat flows into the gas? (b) Ice enclosed in a container does 2.50×10^4 J of work on its container as 7.00×10^3 J of heat flows out of the ice. What is the internal energy change of the ice? (c) The internal energy of a metal block increases by 4.25×10^3 J as 2.40×10^3 J of heat flows into the block. How much work did the block do on its surroundings?

VP19.5.2 How much work does a quantity of gas (volume V, pressure p) do in each of these processes? (a) V increases from 2.00×10^{-3} m^3 to 4.50×10^{-3} m^3 while p is a constant 6.20×10^4 Pa. (b) V is a constant 2.00×10^{-3} m^3 while p increases from 6.00×10^4 Pa to 9.00×10^4 Pa. (c) V decreases from 6.00×10^{-3} m^3 to 3.00×10^{-3} m^3 while p is a constant 1.25×10^5 Pa. (d) V increases from 2.00×10^{-3} m^3 to 4.50×10^{-3} m^3 while p increases from 1.50×10^5 Pa to 5.50×10^5 Pa. The graph of this process on a pV-diagram is a straight line.

VP19.5.3 A quantity of gas is enclosed in a cylinder that has a movable piston. The gas has volume $V = 3.60 \times 10^{-3}$ m^3 and pressure $p = 2.00 \times 10^5$ Pa. (a) First V decreases to 2.40×10^{-3} m^3 while p remains constant; then V remains constant while p increases to 6.00×10^5 Pa. During this process 1.56×10^3 J of heat flows into the gas. For this process, how much work does the gas do, and what is its internal energy change? (b) The gas returns to the same initial volume and pressure. Now p increases to 6.00×10^5 Pa while V remains constant; then V decreases to 2.40×10^{-3} m^3 while p remains constant. For this process, how much work does the gas do, and how much heat flows into the gas?

VP19.5.4 At an atmospheric pressure of 1.01×10^5 Pa, carbon dioxide (CO_2) sublimes at a temperature of −78.5°C. The heat of sublimation is 5.71×10^5 J/kg. You let heat flow into 6.0 kg of solid CO_2 (volume 4.0×10^{-3} m^3) at atmospheric pressure and a constant −78.5°C until it has completely sublimed to gas at −78.5°C. The CO_2 gas occupies a volume of 3.4 m^3. For this process, calculate (a) how much heat flows into the CO_2, (b) how much work the CO_2 does, and (c) the internal energy change of the CO_2.

Be sure to review EXAMPLE 19.6 (Section 19.7) before attempting these problems.

VP19.6.1 For neon gas, $C_V = \frac{3}{2}R$ and $C_p = \frac{5}{2}R$. Calculate the change in internal energy of 4.00 mol of neon if its temperature (a) increases from 20.0°C to 40.0°C at constant volume; (b) increases from 15.0°C to 30.0°C at constant pressure; (c) increases from 12.0°C to 24.0°C while both the pressure and volume increase.

VP19.6.2 For nitrogen gas, $C_V = \frac{5}{2}R$ and $C_p = \frac{7}{2}R$. You have 2.10 mol of nitrogen gas at temperature 27.0°C and pressure $p = 1.00 \times 10^5$ Pa. (a) Calculate the initial volume V of the gas. (b) Calculate the final temperature (in °C) of the gas and its internal energy change if (i) p doubles while V remains constant, (ii) V doubles while p remains constant, and (iii) both p and V double.

VP19.6.3 A quantity of argon gas ($C_V = \frac{3}{2}R$ and $C_p = \frac{5}{2}R$) is at pressure $p = 1.20 \times 10^5$ Pa and occupies volume $V = 0.250$ m^3. Calculate the change in internal energy of the gas if the final pressure and volume of the gas are (a) $p = 2.40 \times 10^5$ Pa, $V = 0.250$ m^3; (b) $p = 1.20 \times 10^5$ Pa, $V = 0.125$ m^3; (c) $p = 1.80 \times 10^5$ Pa, $V = 0.600$ m^3.

VP19.6.4 A quantity of n moles of oxygen gas ($C_V = \frac{5}{2}R$ and $C_p = \frac{7}{2}R$) is at absolute temperature T. You increase the absolute temperature to $2T$. Find the change in internal energy of the gas, the heat flow into the gas, and the work done by the gas if the process you used to increase the temperature is (a) isochoric, (b) isobaric, or (c) adiabatic.

Be sure to review EXAMPLE 19.7 (Section 19.8) before attempting these problems.

VP19.7.1 A quantity of argon gas ($C_V = \frac{3}{2}R$ and $C_p = \frac{5}{2}R$) is at pressure 4.00×10^5 Pa and occupies volume 2.00×10^{-3} m^3. Then the gas expands adiabatically to a new volume 6.00×10^{-3} m^3. Calculate (a) the value of γ for argon, (b) the final pressure of the gas, and (c) the work done by the gas as it expands.

VP19.7.2 A quantity of oxygen gas ($C_V = \frac{5}{2}R$ and $C_p = \frac{7}{2}R$) is at absolute temperature 325 K, is under pressure 1.00×10^5 Pa, and occupies volume 6.50×10^{-3} m^3. After you compress this gas adiabatically, its absolute temperature is 855 K. Calculate (a) the value of γ for oxygen, (b) the final volume of the gas, (c) the final pressure of the gas, and (d) the work done by the gas as it is compressed.

VP19.7.3 Initially 5.00 mol of neon gas ($C_V = \frac{3}{2}R$ and $\gamma = \frac{5}{3}R$) are at absolute temperature 305 K and occupy volume 4.00×10^{-2} m^3. Then the gas expands adiabatically to a new volume of 9.00×10^{-2} m^3. Calculate (a) the initial pressure of the gas, (b) the final pressure of the gas, (c) the final temperature of the gas, and (d) the work done by the gas as it expands.

VP19.7.4 You have 1.25 mol of hydrogen gas ($C_V = \frac{5}{2}R$ and $\gamma = \frac{7}{5}R$) at absolute temperature 325 K. You allow the gas to expand adiabatically to a final temperature of 195 K. (a) How much work does the gas do while being compressed? (b) What is the ratio of its final volume to its initial volume? (c) What is the ratio of the final gas pressure to the initial gas pressure?

BRIDGING PROBLEM Work Done by a van der Waals Gas

The van der Waals equation of state, an approximate representation of the behavior of gases at high pressure, is given by Eq. (18.7): $[p + (an^2/V^2)](V - nb) = nRT$, where a and b are constants having different values for different gases. (In the special case of $a = b = 0$, this is the ideal-gas equation.) (a) Calculate the work done by a gas with this equation of state in an isothermal expansion from V_1 to V_2. (b) For ethane gas (C_2H_6), $a = 0.554$ J·m^3/mol^2 and $b = 6.38 \times 10^{-5}$ m^3/mol. Calculate the work W done by 1.80 mol of ethane when it expands from 2.00×10^{-3} m^3 to 4.00×10^{-3} m^3 at a constant temperature of 300 K. Do the calculation by using (i) the van der Waals equation of state and (ii) the ideal-gas equation of state. (c) For which equation of state is W larger? Why should this be so?

SOLUTION GUIDE

IDENTIFY and SET UP

1. Review the discussion of the van der Waals equation of state in Section 18.1. What is the significance of the quantities a and b?

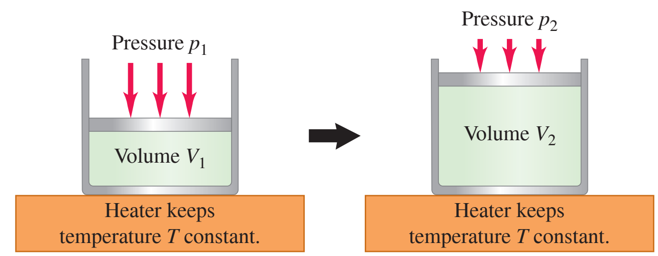
Continued

2. Decide how to find the work done by an expanding gas whose pressure p does not depend on V in the same way as for an ideal gas. (*Hint:* See Section 19.2.)
3. How will you find the work done by an expanding ideal gas?

EXECUTE

4. Find the general expression for the work done by a van der Waals gas as it expands from volume V_1 to volume V_2 (**Fig. 19.22**). (*Hint:* If you set $a = b = 0$ in your result, it should reduce to the expression for the work done by an expanding ideal gas.)
5. Use your result from step 4 to solve part (b) for ethane treated as a van der Waals gas.
6. Use the formula you chose in step 3 to solve part (b) for ethane treated as an ideal gas.

Figure **19.22** A gas undergoes an isothermal expansion.



EVALUATE

7. Is the difference between W for the two equations of state large enough to be significant?
8. Does the term with a in the van der Waals equation of state increase or decrease the amount of work done? What about the term with b? Which one is more important for the ethane in this problem?

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q19.1 For the following processes, is the work done by the system (defined as the expanding or contracting gas) on the environment positive or negative? (a) expansion of the burned gasoline–air mixture in the cylinder of an automobile engine; (b) opening a bottle of champagne; (c) filling a scuba tank with compressed air; (d) partial crumpling of a sealed, empty water bottle as you drive from the mountains down to sea level.

Q19.2 It is not correct to say that an object contains a certain amount of heat, yet an object can transfer heat to another object. How can an object give away something it does not have in the first place?

Q19.3 In which situation must you do more work: inflating a balloon at sea level or inflating the same balloon to the same volume at the summit of Mt. McKinley? Explain in terms of pressure and volume change.

Q19.4 If you are told the initial and final states of a system and the associated change in internal energy, can you determine whether the internal energy change was due to work or to heat transfer? Explain.

Q19.5 Discuss the application of the first law of thermodynamics to a mountaineer who eats food, gets warm and perspires a lot during a climb, and does a lot of mechanical work in raising herself to the summit. The mountaineer also gets warm during the descent. Is the source of this energy the same as the source during the ascent?

Q19.6 When ice melts at 0°C, its volume decreases. Is the internal energy change greater than, less than, or equal to the heat added? How can you tell?

Q19.7 You hold an inflated balloon over a hot-air vent in your house and watch it slowly expand. You then remove it and let it cool back to room temperature. During the expansion, which was larger: the heat added to the balloon or the work done by the air inside it? Explain. (Assume that air is an ideal gas.) Once the balloon has returned to room temperature, how does the net heat gained or lost by the air inside it compare to the net work done on or by the surrounding air?

Q19.8 You bake chocolate chip cookies and put them, still warm, in a container with a loose (not airtight) lid. What kind of process does the air inside the container undergo as the cookies gradually cool to room temperature (isothermal, isochoric, adiabatic, isobaric, or some combination)? Explain.

Q19.9 Imagine a gas made up entirely of negatively charged electrons. Like charges repel, so the electrons exert repulsive forces on each other. Would you expect that the temperature of such a gas would rise, fall, or stay the same in a free expansion? Why?

Q19.10 In an adiabatic process for an ideal gas, the pressure decreases. In this process does the internal energy of the gas increase or decrease? Explain.

Q19.11 When you blow on the back of your hand with your mouth wide open, your breath feels warm. But if you partially close your mouth to form an "o" and then blow on your hand, your breath feels cool. Why?

Q19.12 An ideal gas expands while the pressure is kept constant. During this process, does heat flow into the gas or out of the gas? Justify your answer.

Q19.13 A liquid is irregularly stirred in a well-insulated container and thereby undergoes a rise in temperature. Regard the liquid as the system. Has heat been transferred? How can you tell? Has work been done? How can you tell? Why is it important that the stirring is irregular? What is the sign of ΔU? How can you tell?

Q19.14 When you use a hand pump to inflate the tires of your bicycle, the pump gets warm after a while. Why? What happens to the temperature of the air in the pump as you compress it? Why does this happen? When you raise the pump handle to draw outside air into the pump, what happens to the temperature of the air taken in? Again, why does this happen?

Q19.15 In the carburetor of an aircraft or automobile engine, air flows through a relatively small aperture and then expands. In cool, foggy weather, ice sometimes forms in this aperture even though the outside air temperature is above freezing. Why?

Q19.16 On a sunny day, large "bubbles" of air form on the sunwarmed earth, gradually expand, and finally break free to rise through the atmosphere. Soaring birds and glider pilots are fond of using these "thermals" to gain altitude easily. This expansion is essentially an adiabatic process. Why?

Q19.17 The prevailing winds on the Hawaiian island of Kauai blow from the northeast. The winds cool as they go up the slope of Mt. Waialeale (elevation 1523 m), causing water vapor to condense and rain to fall. There is much more precipitation at the summit than at the base of the mountain. In fact, Mt. Waialeale is the rainiest spot on earth, averaging 11.7 m of rainfall a year. But what makes the winds cool?

Q19.18 Applying the same considerations as in Question Q19.17, explain why the island of Niihau, a few kilometers to the southwest of Kauai, is almost a desert and farms there need to be irrigated.

Q19.19 In a constant-volume process, $dU = nC_V\,dT$. But in a constant-pressure process, it is *not* true that $dU = nC_p\,dT$. Why not?

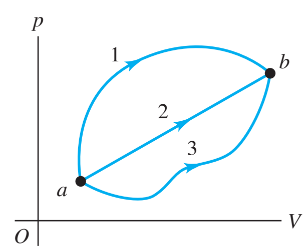
Q19.20 When a gas surrounded by air is compressed adiabatically, its temperature rises even though there is no heat input to the gas. Where does the energy come from to raise the temperature?

Q19.21 When a gas expands adiabatically, it does work on its surroundings. But if there is no heat input to the gas, where does the energy come from to do the work?

Q19.22 The gas used in separating the two uranium isotopes ^{235}U and ^{238}U has the formula UF_6. If you added heat at equal rates to a mole of UF_6 gas and a mole of H_2 gas, which one's temperature would you expect to rise faster? Explain.

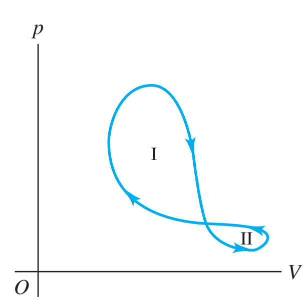
Q19.23 A system is taken from state a to state b along the three paths shown in **Fig. Q19.23**. (a) Along which path is the work done by the system the greatest? The least? (b) If $U_b > U_a$, along which path is the absolute value of the heat transfer, $|Q|$, the greatest? For this path, is heat absorbed or liberated by the system? Explain.

Figure **Q19.23**



Q19.24 A thermodynamic system undergoes a cyclic process as shown in **Fig. Q19.24**. The cycle consists of two closed loops: I and II. (a) Over one complete cycle, does the system do positive or negative work? (b) In each loop, is the net work done by the system positive or negative? (c) Over one complete cycle, does heat flow into or out of the system? (d) In each loop, does heat flow into or out of the system? Explain.

Figure **Q19.24**



EXERCISES

Section 19.2 Work Done During Volume Changes and Section 19.3 Paths Between Thermodynamic States

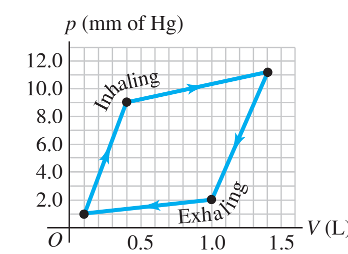
19.1 •• Two moles of an ideal gas are heated at constant pressure from $T = 27°C$ to $T = 107°C$. (a) Draw a pV-diagram for this process. (b) Calculate the work done by the gas.

19.2 • Six moles of an ideal gas are in a cylinder fitted at one end with a movable piston. The initial temperature of the gas is 27.0°C and the pressure is constant. As part of a machine design project, calculate the final temperature of the gas after it has done 2.40×10^3 J of work.

19.3 •• CALC Two moles of an ideal gas are compressed in a cylinder at a constant temperature of 65.0°C until the original pressure has tripled. (a) Sketch a pV-diagram for this process. (b) Calculate the amount of work done.

19.4 •• BIO **Work Done by the Lungs.** The graph in **Fig. E19.4** shows a pV-diagram of the air in a human lung when a person is inhaling and then exhaling a deep breath. Such graphs, obtained in clinical practice, are normally somewhat curved, but we have modeled one as a set of straight lines of the same general shape. (*Important:* The pressure shown is the *gauge* pressure, *not* the absolute pressure.) (a) How many joules of *net* work does this person's lung do during one complete breath? (b) The process illustrated here is somewhat different from those we have been studying, because the pressure change is due to changes in the amount of gas in the lung, not to temperature changes. (Think of your own breathing. Your lungs do not expand because they've gotten hot.) If the temperature of the air in the lung remains a reasonable 20°C, what is the maximum number of moles in this person's lung during a breath?

Figure **E19.4**



19.5 •• CALC During the time 0.305 mol of an ideal gas undergoes an isothermal compression at 22.0°C, 392 J of work is done on it by the surroundings. (a) If the final pressure is 1.76 atm, what was the initial pressure? (b) Sketch a pV-diagram for the process.

19.6 •• A gas undergoes two processes. In the first, the volume remains constant at 0.200 m^3 and the pressure increases from 2.00×10^5 Pa to 5.00×10^5 Pa. The second process is a compression to a volume of 0.120 m^3 at a constant pressure of 5.00×10^5 Pa. (a) In a pV-diagram, show both processes. (b) Find the total work done by the gas during both processes.

19.7 • **Work Done in a Cyclic Process.** (a) In Fig. 19.7a, consider the closed loop $1 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 1$. This is a *cyclic* process in which the initial and final states are the same. Find the total work done by the system in this cyclic process, and show that it is equal to the area enclosed by the loop. (b) How is the work done for the process in part (a) related to the work done if the loop is traversed in the opposite direction, $1 \rightarrow 4 \rightarrow 2 \rightarrow 3 \rightarrow 1$? Explain.

Section 19.4 Internal Energy and the First Law of Thermodynamics

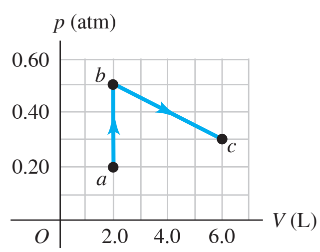
19.8 •• CALC An ideal gas undergoes a process during which the pressure is kept directly proportional to the volume, so that $p = \alpha V$, where α is a positive constant. If the volume changes from V_1 to V_2, how much work is done by the gas? Express your answer in terms of V_1, V_2, and α.

19.9 • A gas in a cylinder expands from a volume of 0.110 m^3 to 0.320 m^3. Heat flows into the gas just rapidly enough to keep the pressure constant at 1.65×10^5 Pa during the expansion. The total heat added is 1.15×10^5 J. (a) Find the work done by the gas. (b) Find the change in internal energy of the gas. (c) Does it matter whether the gas is ideal? Why or why not?

19.10 •• Five moles of an ideal monatomic gas with an initial temperature of 127°C expand and, in the process, absorb 1500 J of heat and do 2100 J of work. What is the final temperature of the gas?

19.11 •• The process *abc* shown in the *pV*-diagram in **Fig. E19.11** involves 0.0175 mol of an ideal gas. (a) What was the lowest temperature the gas reached in this process? Where did it occur? (b) How much work was done by or on the gas from *a* to *b*? From *b* to *c*? (c) If 215 J of heat was put into the gas during *abc*, how many of those joules went into internal energy?

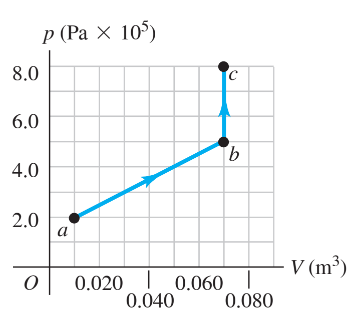
Figure **E19.11**



19.12 • A gas in a cylinder is held at a constant pressure of 1.80×10^5 Pa and is cooled and compressed from 1.70 m^3 to 1.20 m^3. The internal energy of the gas decreases by 1.40×10^5 J. (a) Find the work done by the gas. (b) Find the absolute value of the heat flow, $|Q|$, into or out of the gas, and state the direction of the heat flow. (c) Does it matter whether the gas is ideal? Why or why not?

19.13 •• The *pV*-diagram in **Fig. E19.13** shows a process *abc* involving 0.450 mol of an ideal gas. (a) What was the temperature of this gas at points *a*, *b*, and *c*? (b) How much work was done by or on the gas in this process? (c) How much heat had to be added during the process to increase the internal energy of the gas by 15,000 J?

Figure **E19.13**



19.14 • **Boiling Water at High Pressure.** When water is boiled at a pressure of 2.00 atm, the heat of vaporization is 2.20×10^6 J/kg and the boiling point is 120°C. At this pressure, 1.00 kg of water has a volume of 1.00×10^{-3} m^3, and 1.00 kg of steam has a volume of 0.824 m^3. (a) Compute the work done when 1.00 kg of steam is formed at this temperature. (b) Compute the increase in internal energy of the water.

19.15 • A thermodynamic system undergoes processes in which $|Q| = 100$ J and $|W| = 300$ J. Find Q and W, including whether the quantity is positive or negative, if the change in internal energy is (a) +400 J, (b) +200 J, (c) −200 J, and (d) −400 J.

Section 19.5 Kinds of Thermodynamic Processes, Section 19.6 Internal Energy of an Ideal Gas, and Section 19.7 Heat Capacities of an Ideal Gas

19.16 • During an isothermal compression of an ideal gas, 410 J of heat must be removed from the gas to maintain constant temperature. How much work is done by the gas during the process?

19.17 • A cylinder contains 0.250 mol of carbon dioxide (CO_2) gas at a temperature of 27.0°C. The cylinder is provided with a frictionless piston, which maintains a constant pressure of 1.00 atm on the gas. The gas is heated until its temperature increases to 127.0°C. Assume that the CO_2 may be treated as an ideal gas. (a) Draw a *pV*-diagram for this process. (b) How much work is done by the gas in this process? (c) On what is this work done? (d) What is the change in internal energy of the gas? (e) How much heat was supplied to the gas? (f) How much work would have been done if the pressure had been 0.50 atm?

19.18 • A cylinder contains 0.0100 mol of helium at $T = 27.0$°C. (a) How much heat is needed to raise the temperature to 67.0°C while keeping the volume constant? Draw a *pV*-diagram for this process. (b) If instead the pressure of the helium is kept constant, how much heat is needed to raise the temperature from 27.0°C to 67.0°C? Draw a *pV*-diagram for this process. (c) What accounts for the difference between your answers to parts (a) and (b)? In which case is more heat required? What becomes of the additional heat? (d) If the gas is ideal, what is the change in its internal energy in part (a)? In part (b)? How do the two answers compare? Why?

19.19 • In an experiment to simulate conditions inside an automobile engine, 0.185 mol of air at 780 K and 3.00×10^6 Pa is contained in a cylinder of volume 40.0 cm^3. Then 645 J of heat is transferred to the cylinder. (a) If the volume of the cylinder is constant while the heat is added, what is the final temperature of the air? Assume that the air is essentially nitrogen gas, and use the data in Table 19.1 even though the pressure is not low. Draw a *pV*-diagram for this process. (b) If instead the volume of the cylinder is allowed to increase while the pressure remains constant, repeat part (a).

19.20 •• A quantity of 5.00 mol of a monatomic ideal gas has an initial pressure of $p_a = 1.00 \times 10^5$ Pa and an initial volume of $V_a = 0.0200$ m^3.Two paths are used to take the gas to a final state for which $p_b = 1.80 \times 10^5$ Pa and $V_b = 0.0280$ m^3. (a) Along path 1 the pressure is increased to $p_a = 1.80 \times 10^5$ Pa while the volume is kept constant with a value of 0.0200 m^3. Then the volume is increased to 0.0280 m^3 while the pressure is kept at a constant 1.80×10^5 Pa. Calculate the total Q and total W for this path. (b) Along path 2 the volume is increased to 0.0280 m^3 while the pressure remains 1.00×10^5 Pa and then the pressure is increased to 1.80×10^5 Pa while the volume remains constant at 0.0280 m^3. Calculate the total Q and total W for this path. (c) Calculate the quantity $Q - W$ for each path. Your results should show that for these processes Q and W are path dependent but $Q - W$ is path independent.

19.21 • Heat Q flows into a monatomic ideal gas, and the volume increases while the pressure is kept constant. What fraction of the heat energy is used to do the expansion work of the gas?

19.22 • Three moles of an ideal monatomic gas expands at a constant pressure of 2.50 atm; the volume of the gas changes from 3.20×10^{-2} m^3 to 4.50×10^{-2} m^3. Calculate (a) the initial and final temperatures of the gas; (b) the amount of work the gas does in expanding; (c) the amount of heat added to the gas; (d) the change in internal energy of the gas.

19.23 • An experimenter adds 970 J of heat to 1.75 mol of an ideal gas to heat it from 10.0°C to 25.0°C at constant pressure. The gas does +223 J of work during the expansion. (a) Calculate the change in internal energy of the gas. (b) Calculate γ for the gas.

19.24 • A quantity of 2.00 mol of a monatomic ideal gas undergoes a compression during which the volume decreases from 0.0800 m^3 to 0.0500 m^3 while the pressure stays constant at a value of 1.80×10^4 Pa. (a) What is the work W? Is work done by the gas or on the gas? (b) What is the heat flow Q? Does heat enter or leave the gas? (c) What is the internal energy change for the gas? Does the internal energy of the gas increase or decrease?

19.25 • CALC The temperature of 0.150 mol of an ideal gas is held constant at 77.0°C while its volume is reduced to 25.0% of its initial volume. The initial pressure of the gas is 1.25 atm. (a) Determine the work done by the gas. (b) What is the change in its internal energy? (c) Does the gas exchange heat with its surroundings? If so, how much? Does the gas absorb or liberate heat?

19.26 •• For an ideal gas C_V and C_p are different because of the work W associated with the volume change for a constant-pressure process. To explore the difference between C_V and C_p for a liquid or a solid, consider the process in which 5.00 mol of ethanol is warmed from 10.0°C to 60.0°C while the applied pressure remains a constant 1.00 atm. The molar mass of ethanol is $M = 46.1$ g/mol. (a) What is the mass of 5.00 mol of ethanol? (b) How much heat Q enters the ethanol? (c) How much work W is done by the ethanol because of its thermal expansion? (d) What is the ratio W/Q?

Section 19.8 Adiabatic Processes for an Ideal Gas

19.27 • A monatomic ideal gas that is initially at 1.50×10^5 Pa and has a volume of 0.0800 m^3 is compressed adiabatically to a volume of 0.0400 m^3. (a) What is the final pressure? (b) How much work is done by the gas? (c) What is the ratio of the final temperature of the gas to its initial temperature? Is the gas heated or cooled by this compression?

19.28 •• Five moles of monatomic ideal gas have initial pressure 2.50×10^3 Pa and initial volume 2.10 m^3. While undergoing an adiabatic expansion, the gas does 1480 J of work. What is the final pressure of the gas after the expansion?

19.29 • During an adiabatic expansion the temperature of 0.450 mol of argon (Ar) drops from 66.0°C to 10.0°C. The argon may be treated as an ideal gas. (a) Draw a pV-diagram for this process. (b) How much work does the gas do? (c) What is the change in internal energy of the gas?

19.30 •• A player bounces a basketball on the floor, compressing it to 80.0% of its original volume. The air (assume it is essentially N_2 gas) inside the ball is originally at 20.0°C and 2.00 atm. The ball's inside diameter is 23.9 cm. (a) What temperature does the air in the ball reach at its maximum compression? Assume the compression is adiabatic and treat the gas as ideal. (b) By how much does the internal energy of the air change between the ball's original state and its maximum compression?

19.31 •• On a warm summer day, a large mass of air (atmospheric pressure 1.01×10^5 Pa) is heated by the ground to 26.0°C and then begins to rise through the cooler surrounding air. (This can be treated approximately as an adiabatic process; why?) Calculate the temperature of the air mass when it has risen to a level at which atmospheric pressure is only 0.850×10^5 Pa. Assume that air is an ideal gas, with $\gamma = 1.40$. (This rate of cooling for dry, rising air, corresponding to roughly 1 C° per 100 m of altitude, is called the *dry adiabatic lapse rate.*)

19.32 • A cylinder contains 0.100 mol of an ideal monatomic gas. Initially the gas is at 1.00×10^5 Pa and occupies a volume of 2.50×10^{-3} m^3. (a) Find the initial temperature of the gas in kelvins. (b) If the gas is allowed to expand to twice the initial volume, find the final temperature (in kelvins) and pressure of the gas if the expansion is (i) isothermal; (ii) isobaric; (iii) adiabatic.

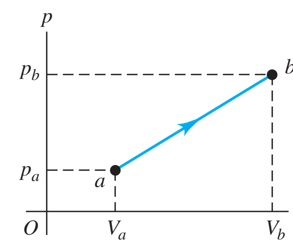
19.33 •• Helium gas undergoes an adiabatic process in which the Kelvin temperature doubles. By what factor does the pressure change?

19.34 • For a diatomic ideal gas, 3.01×10^{20} molecules undergo an adiabatic process in which the temperature increases from 25.0°C to 60.0°C. What is the change in the internal energy of the gas?

PROBLEMS

19.35 •• A quantity of air is taken from state a to state b along a path that is a straight line in the pV-diagram (**Fig. P19.35**). (a) In this process, does the temperature of the gas increase, decrease, or stay the same? Explain. (b) If $V_a = 0.0700$ m^3, $V_b = 0.1100$ m^3, $p_a = 1.00 \times 10^5$ Pa, and $p_b = 1.40 \times 10^5$ Pa, what is the work W done by the gas in this process? Assume that the gas may be treated as ideal.

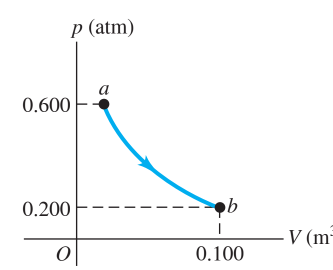
Figure P19.35



19.36 •• When water boils in a pot with a loose lid, the increased pressure forces the lid upward, releasing steam and reducing the pressure in the pot so that the lid drops back down. If the water continues to boil, the lid can rattle upward and downward. Consider a 1 L pot half filled with boiling water. In this case there is half a liter, or $\frac{1}{11.2}$ mol, of air and steam above the water. (a) Estimate the mass of the lid. (b) When the water is rapidly boiling, estimate the height to which the lid jumps during a steam-release event. (c) By multiplying the weight of the lid by the height of the jump, estimate the work done by the steam as the lid moves upward. (d) Treat the process as adiabatic and use the first law of thermodynamics to estimate the drop in the temperature of the steam as it pushes the lid upward.

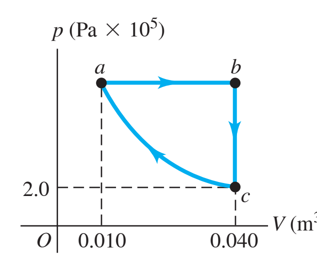
19.37 •• Figure P19.37 shows the pV-diagram for a process in which the temperature of the ideal gas remains constant at 85°C. (a) How many moles of gas are involved? (b) What volume does this gas occupy at a? (c) How much work was done by or on the gas from a to b? (d) By how much did the internal energy of the gas change during this process?

Figure P19.37



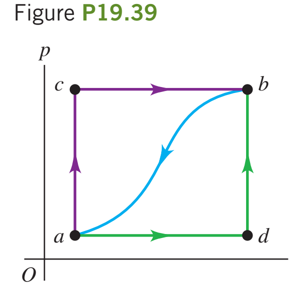
19.38 •• The graph in **Fig. P19.38** shows a pV-diagram for 3.25 mol of *ideal* helium (He) gas. Part ca of this process is isothermal. (a) Find the pressure of the He at point a. (b) Find the temperature of the He at points a, b, and c. (c) How much heat entered or left the He during segments ab, bc, and ca? In each segment, did the heat enter or leave? (d) By how much did the internal energy of the He change from a to b, from b to c, and from c to a? Indicate whether this energy increased or decreased.

Figure P19.38



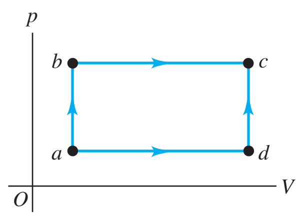
19.39 •• When a system is taken from state a to state b in **Fig. P19.39** along path acb, 90.0 J of heat flows into the system and 60.0 J of work is done by the system. (a) How much heat flows into the system along path adb if the work done by the system is 15.0 J? (b) When the system is returned from b to a along the curved path, the absolute value of the work done by the system is 35.0 J. Does the system absorb or liberate heat? How much heat? (c) If $U_a = 0$ and $U_d = 8.0$ J, find the heat absorbed in the processes ad and db.

Figure P19.39



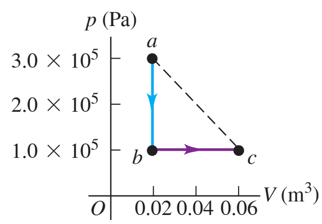
19.40 • A thermodynamic system is taken from state a to state c in **Fig. P19.40** along either path abc or path adc. Along path abc, the work W done by the system is 450 J. Along path adc, W is 120 J. The internal energies of each of the four states shown in the figure are $U_a = 150$ J, $U_b = 240$ J, $U_c = 680$ J, and $U_d = 330$ J. Calculate the heat flow Q for each of the four processes ab, bc, ad, and dc. In each process, does the system absorb or liberate heat?

Figure P19.40



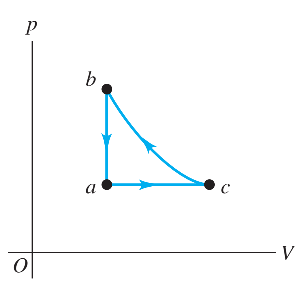
19.41 • A volume of air (assumed to be an ideal gas) is first cooled without changing its volume and then expanded without changing its pressure, as shown by path abc in **Fig. P19.41**. (a) How does the final temperature of the gas compare with its initial temperature? (b) How much heat does the air exchange with its surroundings during process abc? Does the air absorb heat or release heat during this process? Explain. (c) If the air instead expands from state a to state c by the straight-line path shown, how much heat does it exchange with its surroundings?

Figure P19.41



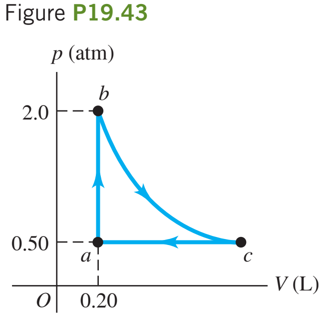
19.42 •• Three moles of an ideal gas are taken around cycle acb shown in **Fig. P19.42**. For this gas, $C_p = 29.1$ J/mol·K. Process ac is at constant pressure, process ba is at constant volume, and process cb is adiabatic. The temperatures of the gas in states a, c, and b are $T_a = 300$ K, $T_c = 492$ K, and $T_b = 600$ K. Calculate the total work W for the cycle.

Figure P19.42



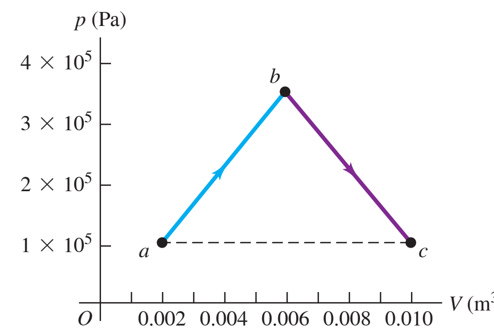
19.43 •• **Figure P19.43** shows a pV-diagram for 0.0040 mol of *ideal* H_2 gas. The temperature of the gas does not change during segment bc. (a) What volume does this gas occupy at point c? (b) Find the temperature of the gas at points a, b, and c. (c) How much heat went into or out of the gas during segments ab, ca, and bc? Indicate whether the heat has gone into or out of the gas. (d) Find the change in the internal energy of this hydrogen during segments ab, bc, and ca. Indicate whether the internal energy increased or decreased during each segment.

Figure P19.43



19.44 •• (a) One-third of a mole of He gas is taken along the path abc shown in **Fig. P19.44**. Assume that the gas may be treated as ideal. How much heat is transferred into or out of the gas? (b) If the gas instead went directly from state a to state c along the horizontal dashed line in Fig. P19.44, how much heat would be transferred into or out of the gas? (c) How does Q in part (b) compare with Q in part (a)? Explain.

Figure P19.44



19.45 •• **DATA** You place a quantity of gas into a metal cylinder that has a movable piston at one end. No gas leaks out of the cylinder as the piston moves. The external force applied to the piston can be varied to change the gas pressure as you move the piston to change the volume of the gas. A pressure gauge attached to the interior wall of the cylinder measures the gas pressure, and you can calculate the volume of the gas from a measurement of the piston's position in the cylinder.

You start with a pressure of 1.0 atm and a gas volume of 3.0 L. Holding the pressure constant, you increase the volume to 5.0 L. Then, keeping the volume constant at 5.0 L, you increase the pressure to 3.0 atm. Next you decrease the pressure linearly as a function of volume until the volume is 3.0 L and the pressure is 2.0 atm. Finally, you keep the volume constant at 3.0 L and decrease the pressure to 1.0 atm, returning the gas to its initial pressure and volume. The walls of the cylinder are good conductors of heat, and you provide the required heat sources and heat sinks so that the necessary heat flows can occur. At these relatively high pressures, you suspect that the ideal-gas equation will not apply with much accuracy. You don't know what gas is in the cylinder or whether it is monatomic, diatomic, or polyatomic. (a) Plot the cycle in the pV-plane. (b) What is the net heat flow for the gas during this cycle? Is there net heat flow into or out of the gas?

19.46 •• Nitrogen gas in an expandable container is cooled from 50.0°C to 10.0°C with the pressure held constant at 3.00×10^5 Pa. The total heat liberated by the gas is 2.50×10^4 J. Assume that the gas may be treated as ideal. Find (a) the number of moles of gas; (b) the change in internal energy of the gas; (c) the work done by the gas. (d) How much heat would be liberated by the gas for the same temperature change if the volume were constant?

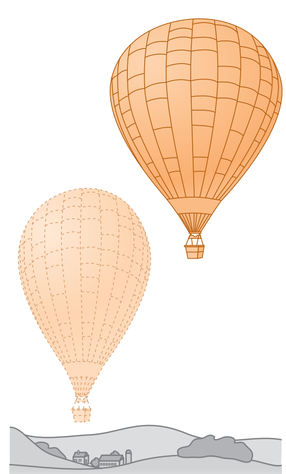
19.47 • **CALC** A cylinder with a frictionless, movable piston like that shown in Fig. 19.5 contains a quantity of helium gas. Initially the gas is at 1.00×10^5 Pa and 300 K and occupies a volume of 1.50 L. The gas then undergoes two processes. In the first, the gas is heated and the piston is allowed to move to keep the temperature at 300 K. This continues until the pressure reaches 2.50×10^4 Pa. In the second process, the gas is compressed at constant pressure until it returns to its original volume of 1.50 L. Assume that the gas may be treated as ideal. (a) In a pV-diagram, show both processes. (b) Find the volume of the gas at the end of the first process, and the pressure and temperature at the end of the second process. (c) Find the total work done by the gas during both processes. (d) What would you have to do to the gas to return it to its original pressure and temperature?

19.48 ••• CP A steel cargo drum has a height of 880 mm and a diameter of 610 mm. With its top removed it has a mass of 17.3 kg. The drum is turned upside down at the surface of the North Atlantic and is pulled downward into the ocean by a robotic submarine. On this day the surface temperature is 23.0°C and the surface air pressure is $p_0 = 101$ kPa. The water temperature decreases linearly with depth to 3.0°C at 1000 m below the surface. As the drum moves downward in the ocean, the air inside the drum is compressed, reducing the upward buoyant force. (a) At what depth y_{neutral} is the barrel neutrally buoyant? (*Hint:* The pressure in the drum is equal to the sea pressure, which at depth y is $p_0 + \rho g y$, where $\rho = 1025\ \text{kg/m}^3$ is the density of seawater. The temperature at depth can be determined using the information above. Together with the ideal-gas law, you can derive a formula for the volume of air at depth y, and therefore a formula for the upward buoyant force as a function of depth.) (b) What is the volume of the air in the drum at depth y_{neutral}?

19.49 ••• Chinook. During certain seasons strong winds called chinooks blow from the west across the eastern slopes of the Rockies and downhill into Denver and nearby areas. Although the mountains are cool, the wind in Denver is very hot; within a few minutes after the chinook wind arrives, the temperature can climb 20 C° ("chinook" refers to a Native American people of the Pacific Northwest). Similar winds occur in the Alps (called foehns) and in southern California (called Santa Anas). (a) Explain why the temperature of the chinook wind rises as it descends the slopes. Why is it important that the wind be fast moving? (b) Suppose a strong wind is blowing toward Denver (elevation 1630 m) from Grays Peak (80 km west of Denver, at an elevation of 4350 m), where the air pressure is 5.60×10^4 Pa and the air temperature is −15.0°C. The temperature and pressure in Denver before the wind arrives are 2.0°C and 8.12×10^4 Pa. By how many Celsius degrees will the temperature in Denver rise when the chinook arrives?

19.50 ••• High-Altitude Research. A large research balloon containing $2.00 \times 10^3\ \text{m}^3$ of helium gas at 1.00 atm and a temperature of 15.0°C rises rapidly from ground level to an altitude at which the atmospheric pressure is only 0.900 atm (**Fig. P19.50**). Assume the helium behaves like an ideal gas and the balloon's ascent is too rapid to permit much heat exchange with the surrounding air. (a) Calculate the volume of the gas at the higher altitude. (b) Calculate the temperature of the gas at the higher altitude. (c) What is the change in internal energy of the helium as the balloon rises to the higher altitude?

Figure **P19.50**



19.51 ••• An air pump has a cylinder 0.250 m long with a movable piston. The pump is used to compress air from the atmosphere (at absolute pressure 1.01×10^5 Pa) into a very large tank at 3.80×10^5 Pa gauge pressure. (For air, $C_V = 20.8\ \text{J/mol} \cdot \text{K}$.) (a) The piston begins the compression stroke at the open end of the cylinder. How far down the length of the cylinder has the piston moved when air first begins to flow from the cylinder into the tank? Assume that the compression is adiabatic. (b) If the air is taken into the pump at 27.0°C, what is the temperature of the compressed air? (c) How much work does the pump do in putting 20.0 mol of air into the tank?

19.52 •• A certain ideal gas has molar heat capacity at constant volume C_V. A sample of this gas initially occupies a volume V_0 at pressure p_0 and absolute temperature T_0. The gas expands isobarically to a volume $2V_0$ and then expands further adiabatically to a final volume $4V_0$. (a) Draw a pV-diagram for this sequence of processes. (b) Compute the total work done by the gas for this sequence of processes. (c) Find the final temperature of the gas. (d) Find the absolute value of the total heat flow $|Q|$ into or out of the gas for this sequence of processes, and state the direction of heat flow.

19.53 • A monatomic ideal gas expands slowly to twice its original volume, doing 450 J of work in the process. Find the heat added to the gas and the change in internal energy of the gas if the process is (a) isothermal; (b) adiabatic; (c) isobaric.

19.54 •• CALC A cylinder with a piston contains 0.250 mol of oxygen at 2.40×10^5 Pa and 355 K. The oxygen may be treated as an ideal gas. The gas first expands isobarically to twice its original volume. It is then compressed isothermally back to its original volume, and finally it is cooled isochorically to its original pressure. (a) Show the series of processes on a pV-diagram. Compute (b) the temperature during the isothermal compression; (c) the maximum pressure; (d) the total work done by the piston on the gas during the series of processes.

19.55 • Use the conditions and processes of Problem 19.54 to compute (a) the work done by the gas, the heat added to it, and its internal energy change during the initial expansion; (b) the work done, the heat added, and the internal energy change during the final cooling; (c) the internal energy change during the isothermal compression.

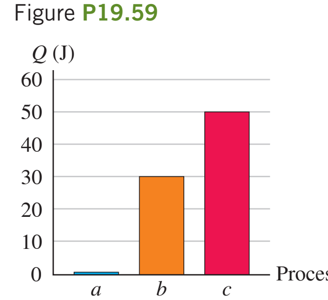
19.56 •• CALC A cylinder with a piston contains 0.150 mol of nitrogen at 1.80×10^5 Pa and 300 K. The nitrogen may be treated as an ideal gas. The gas is first compressed isobarically to half its original volume. It then expands adiabatically back to its original volume, and finally it is heated isochorically to its original pressure. (a) Show the series of processes in a pV-diagram. (b) Compute the temperatures at the beginning and end of the adiabatic expansion. (c) Compute the minimum pressure.

19.57 • Use the conditions and processes of Problem 19.56 to compute (a) the work done by the gas, the heat added to it, and its internal energy change during the initial compression; (b) the work done by the gas, the heat added to it, and its internal energy change during the adiabatic expansion; (c) the work done, the heat added, and the internal energy change during the final heating.

19.58 • Comparing Thermodynamic Processes. In a cylinder, 1.20 mol of an ideal monatomic gas, initially at 3.60×10^5 Pa and 300 K, expands until its volume triples. Compute the work done by the gas if the expansion is (a) isothermal; (b) adiabatic; (c) isobaric. (d) Show each process in a pV-diagram. In which case is the absolute value of the work done by the gas greatest? Least? (e) In which case is the absolute value of the heat transfer greatest? Least? (f) In which case is the absolute value of the change in internal energy of the gas greatest? Least?

19.59 •• DATA You have recorded measurements of the heat flow Q into 0.300 mol of a gas that starts at $T_1 = 20.0$°C and ends at a temperature T_2. You measured Q for three processes: one isobaric, one isochoric, and one adiabatic. In each case, T_2 was the same. **Figure P19.59** summarizes your results. But you lost a page from your lab notebook and don't have a record of the value of T_2; you also don't know which process was isobaric, isochoric, or adiabatic. Each process was done at a sufficiently low pressure for the gas to be treated as ideal. (a) Identify each process a, b, or c as isobaric, isochoric, or adiabatic. (b) What is the value of T_2? (c) How much work is done by the gas in each process? (d) For which process is the magnitude of the volume change the greatest? (e) For each process, does the volume of the gas increase, decrease, or stay the same?

Figure **P19.59**



19.60 •• DATA You compress a gas in an insulated cylinder—no heat flows into or out of the gas. The gas pressure is fairly low, so treating the gas as ideal is a good approximation. When you measure the pressure as a function of the volume of the gas, you obtain these results:

V **(L)**	2.50	2.02	1.48	1.01	0.50
p **(atm)**	0.101	0.139	0.202	0.361	0.952

(a) Graph $\log(p)$ versus $\log(V)$, with p in Pa and V in m^3. Explain why the data points fall close to a straight line. (b) Use your graph to calculate γ for the gas. Is the gas monatomic, diatomic, or polyatomic? (c) When $p = 0.101$ atm and $V = 2.50$ L, the temperature is 22.0°C. Apply the ideal-gas equation and calculate the temperature for each of the other pairs of p and V values. In this compression, does the temperature of the gas increase, decrease, or stay constant?

19.61 ••• A pneumatic shock absorber consists of a cylinder with a radius of 1.50 cm and a length that varies from a maximum of 30.0 cm to a minimum of 30.0 cm$/v$, where v is the compression ratio. For ease of manufacturing, v must be an integer. When the cylinder is fully extended and cooled, the air inside has an ambient pressure of 101 kPa and ambient temperature of 30.0°C. A quick adiabatic compression heats the air. The cylinder then expands isothermally until it reaches its maximum length. At that time the length remains fixed and the cylinder cools isochorically, returning to the ambient temperature. (a) How much work is done by the air during a full cycle, in terms of v? (b) If the air temperature cannot exceed 400°C, what is the maximum integer value of v? (c) If at least 25.0 J of work must be done by the gas in a full-compression cycle, what is the minimum allowable integer value of v? (d) What is the only permissible value of v? (e) How much heat leaves the air in the cylinder during the isochoric process if v has the value calculated in (d)?

CHALLENGE PROBLEM

19.62 •• Engine Turbochargers and Intercoolers. The power output of an automobile engine is directly proportional to the mass of air that can be forced into the volume of the engine's cylinders to react chemically with gasoline. Many cars have a *turbocharger,* which compresses the air before it enters the engine, giving a greater mass of air per volume. This rapid, essentially adiabatic compression also heats the air. To compress it further, the air then passes through an *intercooler* in which the air exchanges heat with its surroundings at essentially constant pressure. The air is then drawn into the cylinders. In a typical installation, air is taken into the turbocharger at atmospheric pressure $(1.01 \times 10^5$ Pa$)$, density $\rho = 1.23$ kg/m^3, and temperature 15.0°C. It is compressed adiabatically to 1.45×10^5 Pa. In the intercooler, the air is cooled to the original temperature of 15.0°C at a constant pressure of 1.45×10^5 Pa. (a) Draw a pV-diagram for this sequence of processes. (b) If the volume of one of the engine's cylinders is 575 cm^3, what mass of air exiting from the intercooler will fill the cylinder at 1.45×10^5 Pa? Compared to the power output of an engine that takes in air at 1.01×10^5 Pa at 15.0°C, what percentage increase in power is obtained by using the turbocharger and intercooler? (c) If the intercooler is not used, what mass of air exiting from the turbocharger will fill the cylinder at 1.45×10^5 Pa? Compared to the power output of an engine that takes in air at 1.01×10^5 Pa at 15.0°C, what percentage increase in power is obtained by using the turbocharger alone?

MCAT-STYLE PASSAGE PROBLEMS

BIO Anesthetic Gases. One type of gas mixture used in anesthesiology is a 50%/50% mixture (by volume) of nitrous oxide (N_2O) and oxygen (O_2), which can be premixed and kept in a cylinder for later use. Because these two gases don't react chemically at or below 2000 psi, at typical room temperatures they form a homogeneous single gas phase, which can be considered an ideal gas. If the temperature drops below −6°C, however, N_2O may begin to condense out of the gas phase. Then any gas removed from the cylinder will initially be nearly pure O_2; as the cylinder empties, the proportion of O_2 will decrease until the gas coming from the cylinder is nearly pure N_2O.

19.63 In a test of the effects of low temperatures on the gas mixture, a cylinder filled at 20.0°C to 2000 psi (gauge pressure) is cooled slowly and the pressure is monitored. What is the expected pressure at −5.00°C if the gas remains a homogeneous mixture? (a) 500 psi; (b) 1500 psi; (c) 1830 psi; (d) 1920 psi.

19.64 In another test, the valve of a 500 L cylinder full of the gas mixture at 2000 psi (gauge pressure) is opened wide so that the gas rushes out of the cylinder very rapidly. Why might some N_2O condense during this process? (a) This is an isochoric process in which the pressure decreases, so the temperature also decreases. (b) Because of the rapid expansion, heat is removed from the system, so the internal energy and temperature of the gas decrease. (c) This is an isobaric process, so as the volume increases, the temperature decreases proportionally. (d) With the rapid expansion, the expanding gas does work with no heat input, so the internal energy and temperature of the gas decrease.

19.65 You have a cylinder that contains 500 L of the gas mixture pressurized to 2000 psi (gauge pressure). A regulator sets the gas flow to deliver 8.2 L/min at atmospheric pressure. Assume that this flow is slow enough that the expansion is isothermal and the gases remain mixed. How much time will it take to empty the cylinder? (a) 1 h; (b) 33 h; (c) 57 h; (d) 140 h.

19.66 In a hospital, pure oxygen may be delivered at 50 psi (gauge pressure) and then mixed with N_2O. What volume of oxygen at 20°C and 50 psi (gauge pressure) should be mixed with 1.7 kg of N_2O to get a 50%/50% mixture by volume at 20°C? (a) 0.21 m^3; (b) 0.27 m^3; (c) 1.9 m^3; (d) 100 m^3.

ANSWERS

Chapter Opening Question ?

(ii) The work done by a gas as its volume changes from V_1 to V_2 is equal to the integral $\int p\,dV$ between those two volume limits. If the volume of the gas contracts, the final volume V_2 is less than the initial volume V_1 and the gas does negative work. Propelling the locomotive requires that the gas do positive work, so the gas doesn't contribute to propulsion while contracting.

Key Example VARIATION Problems

VP19.5.1 (a) 5.75×10^3 J (b) -3.20×10^4 J (c) -1.85×10^3 J (d) 875 J
VP19.5.2 (a) 155 J (b) zero (c) -375 J (d) 875 J
VP19.5.3 (a) $W = -2.40 \times 10^2$ J, $\Delta U = 1.80 \times 10^3$ J (b) $W = -7.20 \times 10^2$ J, $Q = 1.08 \times 10^3$ J
VP19.5.4 (a) 3.4×10^6 J (b) 3.4×10^5 J (c) 3.1×10^6 J
VP19.6.1 (a) 998 J (b) 748 J (c) 599 J
VP19.6.2 (a) 5.24×10^{-2} m^3 (b) (i) 327°C, 1.31×10^4 J (ii) 327°C, 1.31×10^4 J (iii) 927°C, 3.93×10^4 J
VP19.6.3 (a) 4.50×10^4 J (b) -2.25×10^4 J (c) 1.17×10^5 J
VP19.6.4 (a) $\Delta U = \frac{5}{2}nRT$, $Q = \frac{5}{2}nRT$, $W = 0$ (b) $\Delta U = \frac{5}{2}nRT$, $Q = \frac{7}{2}nRT$, $W = nRT$ (c) $\Delta U = \frac{5}{2}nRT$, $Q = 0$, $W = -\frac{5}{2}nRT$
VP19.7.1 (a) $\frac{5}{3}$ (b) 6.41×10^4 Pa (c) 623 J
VP19.7.2 (a) $\frac{7}{5}$ (b) 5.79×10^{-4} m^3 (c) 2.95×10^6 Pa (d) -2.65×10^3 J
VP19.7.3 (a) 3.17×10^5 Pa (b) 8.20×10^4 Pa (c) 178 K (d) 7.94×10^3 J
VP19.7.4 (a) -3.38×10^3 J (b) 3.59 (c) 0.167

Bridging Problem

(a) $W = nRT \ln\left[\dfrac{V_2 - nb}{V_1 - nb}\right] + an^2\left[\dfrac{1}{V_2} - \dfrac{1}{V_1}\right]$

(b) (i) $W = 2.80 \times 10^3$ J, (ii) $W = 3.11 \times 10^3$ J

(c) ideal gas, for which there is no attraction between molecules

? The second law of thermodynamics tells us that heat naturally flows from a hot object (such as molten lava, shown here flowing into the ocean in Hawaii) to a cold one (such as ocean water, which is heated by the lava to make steam). Is it *ever* possible for heat to flow from a cold object to a hot one? (i) Yes, no matter what the temperature difference; (ii) yes, but for only certain temperature differences; (iii) no; (iv) answer depends on the compositions of the two objects.

20 The Second Law of Thermodynamics

LEARNING OUTCOMES

In this chapter, you'll learn...

20.1 The difference between reversible and irreversible processes.
20.2 What a heat engine is, and how to calculate its efficiency.
20.3 The physics of internal-combustion engines.
20.4 How refrigerators and heat engines are related, and how to analyze the performance of a refrigerator.
20.5 How the second law of thermodynamics sets limits on the efficiency of engines and the performance of refrigerators.
20.6 How to do calculations involving the idealized Carnot cycle for engines and refrigerators.
20.7 What is meant by entropy, and how to use this concept to analyze thermodynamic processes.
20.8 How to use the concept of microscopic states to understand entropy.

You'll need to review...

17.3 The Kelvin scale.
18.3 The Boltzmann constant.
19.1–19.8 Thermodynamic processes; first law of thermodynamics; free expansion of a gas.

Many thermodynamic processes proceed naturally in one direction but not the opposite. For example, heat by itself always flows from a hot object to a cooler object, never the reverse. Heat flow from a cool object to a hot object would not violate the first law of thermodynamics; energy would be conserved. But it doesn't happen in nature. Why not? As another example, note that it is easy to convert mechanical energy completely into heat; this happens every time we use a car's brakes to stop it. In the reverse direction, there are plenty of devices that convert heat *partially* into mechanical energy. (An automobile engine is an example.) But no one has ever managed to build a machine that converts heat *completely* into mechanical energy. Again, why not?

The answer to both of these questions has to do with the *directions* of thermodynamic processes and is called the *second law of thermodynamics.* This law places fundamental limitations on the efficiency of an engine or a power plant. It also places limitations on the minimum energy input needed to operate a refrigerator. So the second law is directly relevant for many important practical problems.

We can also state the second law in terms of the concept of *entropy,* a quantitative measure of the degree of randomness of a system. The idea of entropy helps explain why ink mixed with water never spontaneously unmixes and why we never observe a host of other seemingly possible processes.

20.1 DIRECTIONS OF THERMODYNAMIC PROCESSES

Thermodynamic processes that occur in nature are all **irreversible processes.** These are processes that proceed spontaneously in one direction but not the other (**Fig. 20.1a**). The flow of heat from a hot object to a cooler object is irreversible, as is the free expansion of a gas discussed in Sections 19.3 and 19.6. Sliding a book across a table converts mechanical energy into heat by friction; this process is irreversible, for no one has ever observed the reverse process (in which a book initially at rest on the table would spontaneously start moving and the table and book would cool down). Our main topic for this

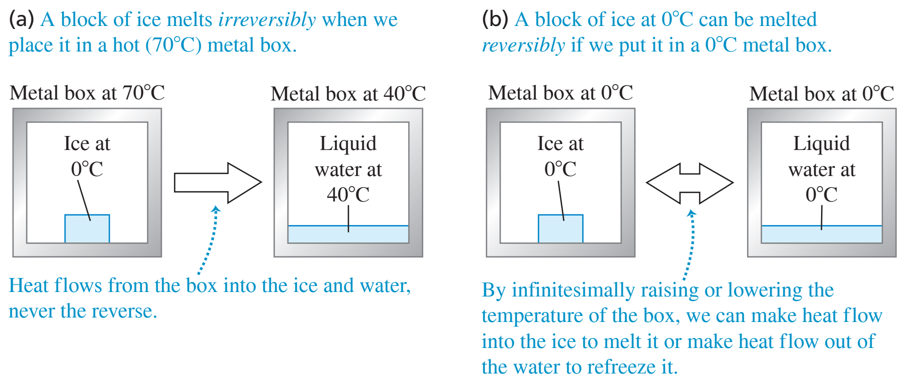


Figure **20.1** Reversible and irreversible processes.

chapter is the *second law of thermodynamics,* which determines the preferred direction for such processes.

Despite this preferred direction for every natural process, we can think of a class of idealized processes that *would* be reversible. A system that undergoes such an idealized **reversible process** is always very close to being in thermodynamic equilibrium within itself and with its surroundings. Any change of state that takes place can then be reversed by making only an infinitesimal change in the conditions of the system. For example, we can reverse heat flow between two objects whose temperatures differ only infinitesimally by making only a very small change in one temperature or the other (Fig. 20.1b).

Reversible processes are thus **equilibrium processes,** with the system always in thermodynamic equilibrium. Of course, if a system were *truly* in thermodynamic equilibrium, no change of state would take place. Heat would not flow into or out of a system with truly uniform temperature throughout, and a system that is truly in mechanical equilibrium would not expand and do work against its surroundings. A reversible process is an idealization that can never be precisely attained in the real world. But by making the temperature gradients and the pressure differences in the substance very small, we can keep the system very close to equilibrium states and make the process nearly reversible.

By contrast, heat flow with finite temperature difference, free expansion of a gas, and conversion of work to heat by friction are all *irreversible* processes; no small change in conditions could make any of them go the other way. They are also all *nonequilibrium* processes, in that the system is not in thermodynamic equilibrium at any point until the end of the process.

Disorder and Thermodynamic Processes

There is a relationship between the direction of a process and the *disorder* or *randomness* of the resulting state. For example, imagine a thousand names written on file cards and arranged in alphabetical order. Throw the alphabetized stack of cards into the air, and they will likely come down in a random, disordered state. In the free expansion of a gas discussed in Sections 19.3 and 19.6, the air is more disordered after it has expanded into the entire box than when it was confined in one side, just as your clothes are more disordered when scattered all over your floor than when confined to your closet.

Similarly, macroscopic kinetic energy is energy associated with organized, coordinated motions of many molecules, but heat transfer involves changes in energy of random, disordered molecular motion. Therefore conversion of mechanical energy into heat involves an increase of randomness or disorder.

In the following sections we'll introduce the second law of thermodynamics by considering two broad classes of devices: *heat engines,* which are partly successful in converting heat into work, and *refrigerators,* which are partly successful in transporting heat from cooler to hotter objects.

TEST YOUR UNDERSTANDING OF SECTION 20.1 Your left and right hands are normally at the same temperature, just like the metal box and ice in Fig. 20.1b. Is rubbing your hands together to warm them (i) a reversible process or (ii) an irreversible process?

ANSWER

(ii) Like sliding a book across a table, rubbing your hands together uses friction to convert mechanical energy into heat. The (impossible) reverse process would involve your hands spontaneously getting colder, with the released energy forcing your hands to move rhythmically back and forth!

20.2 HEAT ENGINES

The essence of our technological society is the ability to use sources of energy other than muscle power. Sometimes, mechanical energy is directly available; water power and wind power are examples. But most of our energy comes from the burning of fossil fuels (coal, oil, and gas) and from nuclear reactions. They supply energy that is transferred as *heat.* This is directly useful for heating buildings, for cooking, and for chemical processing, but to operate a machine or propel a vehicle, we need *mechanical* energy.

Thus it's important to know how to take heat from a source and convert as much of it as possible into mechanical energy or work. This is what happens in gasoline engines in automobiles, jet engines in airplanes, steam turbines in electric power plants, and many other systems. Closely related processes occur in the animal kingdom; food energy is "burned" (that is, carbohydrates combine with oxygen to yield water, carbon dioxide, and energy) and partly converted to mechanical energy as an animal's muscles do work on its surroundings.

Figure 20.2 All motorized vehicles other than purely electric vehicles use heat engines for propulsion. (Hybrid vehicles use their internal-combustion engine to help charge the batteries for the electric motor.)

Any device that transforms heat partly into work or mechanical energy is called a **heat engine** (**Fig. 20.2**). Usually, a quantity of matter inside the engine undergoes inflow and outflow of heat, expansion and compression, and sometimes change of phase. We call this matter the **working substance** of the engine. In internal-combustion engines, such as those used in automobiles, the working substance is a mixture of air and fuel; in a steam turbine it is water.

The simplest kind of engine to analyze is one in which the working substance undergoes a **cyclic process,** a sequence of processes that eventually leaves the substance in the same state in which it started. In a steam turbine the water is recycled and used over and over. Internal-combustion engines do not use the same air over and over, but we can still analyze them in terms of cyclic processes that approximate their actual operation.

Hot and Cold Reservoirs

All heat engines *absorb* heat from a source at a relatively high temperature, perform some mechanical work, and *discard* or *reject* some heat at a lower temperature. As far as the engine is concerned, the discarded heat is wasted. In internal-combustion engines the waste heat is that discarded in the hot exhaust gases and the cooling system; in a steam turbine it is the heat that must flow out of the used steam to condense and recycle the water.

When a system is carried through a cyclic process, its initial and final internal energies are equal, so the first law of thermodynamics requires that

$$U_2 - U_1 = 0 = Q - W \qquad \text{so} \qquad Q = W$$

That is, the net heat flowing into the engine in a cyclic process equals the net work done by the engine.

When we analyze heat engines, it helps to think of two objects with which the working substance of the engine can interact. One of these, called the *hot reservoir,* represents the heat source; it can give the working substance large amounts of heat at a constant temperature T_H without appreciably changing its own temperature. The other object, called the *cold reservoir,* can absorb large amounts of discarded heat from the engine at a constant lower temperature T_C. In a steam-turbine system the flames and hot gases in the boiler are the hot reservoir, and the cold water and air used to condense and cool the used steam are the cold reservoir.

We denote the quantities of heat transferred from the hot and cold reservoirs as Q_H and Q_C, respectively. A quantity of heat Q is positive when heat is transferred *into* the working substance and is negative when heat leaves the working substance. Thus in a heat engine, Q_H is positive but Q_C is negative, representing heat *leaving* the working substance. This sign convention is consistent with the rules we stated in Section 19.1; we'll continue to use those rules here. For clarity, we'll often state the relationships in terms of the absolute values of the Q's and W's because absolute values are always positive.

Energy-Flow Diagrams and Efficiency

We can represent the energy transformations in a heat engine by the *energy-flow diagram* of **Fig. 20.3**. The engine itself is represented by the circle. The amount of heat Q_H supplied to the engine by the hot reservoir is proportional to the width of the incoming "pipeline" at the top of the diagram. The width of the outgoing pipeline at the bottom is proportional to the magnitude $|Q_C|$ of the heat rejected in the exhaust. The branch line to the right represents the portion of the heat supplied that the engine converts to mechanical work, W.

When an engine repeats the same cycle over and over, Q_H and Q_C represent the quantities of heat absorbed and rejected by the engine *during one cycle;* Q_H is positive, and Q_C is negative. The *net* heat Q absorbed per cycle is

$$Q = Q_H + Q_C = |Q_H| - |Q_C| \tag{20.1}$$

The useful output of the engine is the net work W done by the working substance. From the first law,

$$W = Q = Q_H + Q_C = |Q_H| - |Q_C| \tag{20.2}$$

Ideally, we would like to convert *all* the heat Q_H into work; in that case we would have $Q_H = W$ and $Q_C = 0$. Experience shows that this is impossible; there is always some heat wasted, and Q_C *is never zero.* We define the **thermal efficiency** of an engine, denoted by e, as the quotient

$$e = \frac{W}{Q_H} \tag{20.3}$$

The thermal efficiency e represents the fraction of Q_H that *is* converted to work. To put it another way, e is what you get divided by what you pay for. This is always less than unity, an all-too-familiar experience! In terms of the flow diagram of Fig. 20.3, the most efficient engine is one for which the branch pipeline representing the work output is as wide as possible and the exhaust pipeline representing the heat thrown away is as narrow as possible.

When we substitute the two expressions for W given by Eq. (20.2) into Eq. (20.3), we get the following equivalent expressions for e:

Thermal efficiency of an engine

Work done by engine

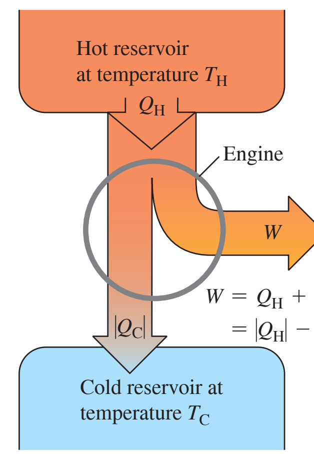
Heat rejected by engine

$$e = \frac{W}{Q_H} = 1 + \frac{Q_C}{Q_H} = 1 - \left|\frac{Q_C}{Q_H}\right| \tag{20.4}$$

Heat absorbed by engine

Note that e is a quotient of two energy quantities and thus is a pure number, without units. Of course, we must always express W, Q_H, and Q_C in the same units.

Figure **20.3** Schematic energy-flow diagram for a heat engine.



BIO APPLICATION Biological Efficiency Although a biological organism is not a heat engine, the concept of efficiency still applies: Here e is the ratio of the work done to the energy that was used to do that work. To exercise on a stationary bike, your body must first convert the chemical-bond energy in glucose to chemical-bond energy in ATP (adenosine triphosphate), then convert energy from ATP into motion of your leg muscles, and finally convert muscular motion into motion of the pedals. The overall efficiency of this entire process is only about 25%. The remaining 75% of the energy liberated from glucose goes into heat.

PROBLEM-SOLVING STRATEGY 20.1 Heat Engines

Problems involving heat engines are, fundamentally, problems in the first law of thermodynamics. You should review Problem-Solving Strategy 19.1 (Section 19.4).

IDENTIFY *the relevant concepts:* A heat engine is any device that converts heat partially to work, as shown schematically in Fig. 20.3. We'll see in Section 20.4 that a refrigerator is essentially a heat engine running in reverse, so many of the same concepts apply.

SET UP *the problem* as suggested in Problem-Solving Strategy 19.1: Use Eq. (20.4) if the thermal efficiency of the engine is relevant. Sketch an energy-flow diagram like Fig. 20.3.

EXECUTE *the solution* as follows:

1. Be careful with the sign conventions for W and the various Q's. W is positive when the system expands and does work; W is negative when the system is compressed and work is done on it. Each Q is positive if it represents heat entering the system and is negative if it represents heat leaving the system. When you know that a quantity is negative, such as Q_C in the above discussion, it sometimes helps to write it as $Q_C = -|Q_C|$.
2. Power is work per unit time ($P = W/t$), and rate of heat transfer (heat current) H is heat transfer per unit time ($H = Q/t$). In problems involving these concepts it helps to ask, "What is W or Q in one second (or one hour)?"
3. Keeping steps 1 and 2 in mind, solve for the target variables.

EVALUATE *your answer:* Use the first law of thermodynamics to check your results. Pay particular attention to algebraic signs.

EXAMPLE 20.1 Analyzing a heat engine

WITH VARIATION PROBLEMS

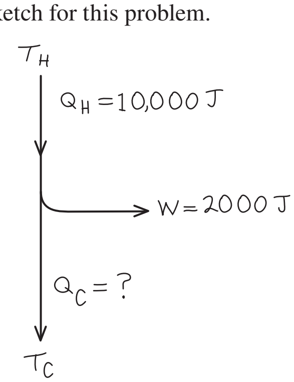
A gasoline truck engine takes in 10,000 J of heat and delivers 2000 J of mechanical work per cycle. The heat is obtained by burning gasoline with heat of combustion $L_c = 5.0 \times 10^4$ J/g. (a) What is the thermal efficiency of this engine? (b) How much heat is discarded in each cycle? (c) If the engine goes through 25 cycles per second, what is its power output in watts? In horsepower? (d) How much gasoline is burned in each cycle? (e) How much gasoline is burned per second? Per hour?

IDENTIFY and SET UP This problem concerns a heat engine, so we can use the ideas of this section. **Figure 20.4** is our energy-flow diagram for one cycle. In each cycle the engine does $W = 2000$ J of work and takes in heat $Q_H = 10{,}000$ J. We use Eq. (20.4), in the form $e = W/Q_H$, to find the thermal efficiency. We use Eq. (20.2) to find the amount of heat Q_C rejected per cycle. The heat of combustion tells us how much gasoline must be burned per cycle and hence per unit time. The power output is the time rate at which the work W is done.

EXECUTE (a) From Eq. (20.4), the thermal efficiency is

$$e = \frac{W}{Q_H} = \frac{2000\text{ J}}{10{,}000\text{ J}} = 0.20 = 20\%$$

Figure 20.4 Our sketch for this problem.



(b) From Eq. (20.2), $W = Q_H + Q_C$, so

$$Q_C = W - Q_H = 2000\text{ J} - 10{,}000\text{ J} = -8000\text{ J}$$

That is, 8000 J of heat leaves the engine during each cycle.

(c) The power P equals the work per cycle multiplied by the number of cycles per second:

$$P = (2000\text{ J/cycle})(25\text{ cycles/s}) = 50{,}000\text{ W} = 50\text{ kW}$$
$$= (50{,}000\text{ W})\frac{1\text{ hp}}{746\text{ W}} = 67\text{ hp}$$

(d) Let m be the mass of gasoline burned during each cycle. Then $Q_H = mL_c$ and

$$m = \frac{Q_H}{L_c} = \frac{10{,}000\text{ J}}{5.0 \times 10^4\text{ J/g}} = 0.20\text{ g}$$

(e) The mass of gasoline burned per second equals the mass per cycle multiplied by the number of cycles per second:

$$(0.20\text{ g/cycle})(25\text{ cycles/s}) = 5.0\text{ g/s}$$

The mass burned per hour is

$$(5.0\text{ g/s})\frac{3600\text{ s}}{1\text{ h}} = 18{,}000\text{ g/h} = 18\text{ kg/h}$$

EVALUATE An efficiency of 20% is fairly typical for cars and trucks if W includes only the work delivered to the wheels. We can check the mass burned per hour by expressing it in miles per gallon ("mileage"). The density of gasoline is about 0.70 g/cm^3, so this is about 25,700 cm^3, 25.7 L, or 6.8 gallons of gasoline per hour. If the truck is traveling at 55 mi/h (88 km/h), this represents fuel consumption of 8.1 miles/gallon (3.4 km/L). This is a fairly typical mileage for large trucks.

KEYCONCEPT A heat engine takes in energy in the form of heat from a hot reservoir, uses some of that energy to do work, and rejects the remaining energy as heat that goes into a cold reservoir. The fraction of the heat from the hot reservoir that is converted into work is called the thermal efficiency of the engine.

TEST YOUR UNDERSTANDING OF SECTION 20.2 Rank the following heat engines in order from highest to lowest thermal efficiency. (i) An engine that in one cycle absorbs 5000 J of heat and rejects 4500 J of heat; (ii) an engine that in one cycle absorbs 25,000 J of heat and does 2000 J of work; (iii) an engine that in one cycle does 400 J of work and rejects 2800 J of heat.

ANSWER

(iii), (i), (ii) From Eq. (20.4) the efficiency is $e = W/Q_H$, and from Eq. (20.2) $W = Q_H + Q_C = |Q_H| - |Q_C|$. For engine (i) $Q_H = 5000$ J and $Q_C = -4500$ J, so $W = 5000\text{ J} + (-4500\text{ J}) = 500$ J and $e = (500\text{ J})/(5000\text{ J}) = 0.100$. For engine (ii) $Q_H = 25{,}000$ J and $W = 2000$ J, so $e = (2000\text{ J})/(25{,}000\text{ J}) = 0.080$. For engine (iii) $W = 400$ J and $Q_C = -2800$ J, so $Q_H = W - Q_C = 400\text{ J} - (-2800\text{ J}) = 3200$ J and $e = (400\text{ J})/(3200\text{ J}) = 0.125$.

20.3 INTERNAL-COMBUSTION ENGINES

The gasoline engine, used in automobiles and many other types of machinery, is a familiar example of a heat engine. Let's look at its thermal efficiency. **Figure 20.5** shows the operation of one type of gasoline engine. First a mixture of air and gasoline vapor flows into a cylinder through an open intake valve while the piston descends, increasing the volume of the cylinder from a minimum of V (when the piston is all the way up) to a maximum of rV (when it is all the way down). The quantity r is called the **compression ratio;** for present-day automobile engines its value is typically 8 to 10. At the end of this *intake stroke,* the intake valve closes and the mixture is compressed, approximately adiabatically, to volume V during the *compression stroke*. The mixture is then ignited by the spark plug, and the heated gas expands, approximately adiabatically, back to volume rV, pushing on the piston and doing work; this is the *power stroke*. Finally, the exhaust valve opens, and the combustion products are pushed out (during the *exhaust stroke*), leaving the cylinder ready for the next intake stroke.

The Otto Cycle

Figure 20.6 (next page) is a pV-diagram for an idealized model of the thermodynamic processes in a gasoline engine. This model is called the **Otto cycle.** At point a the gasoline–air mixture has entered the cylinder. The mixture is compressed adiabatically to point b and is then ignited. Heat Q_H is added to the system by the burning gasoline along line bc, and the power stroke is the adiabatic expansion to d. The gas is cooled to the temperature of the outside air along line da; during this process, heat $|Q_C|$ is rejected. This gas leaves the engine as exhaust and does not enter the engine again. But since an equivalent amount of gasoline and air enters, we may consider the process to be cyclic.

Figure **20.5** Cycle of a four-stroke internal-combustion engine.

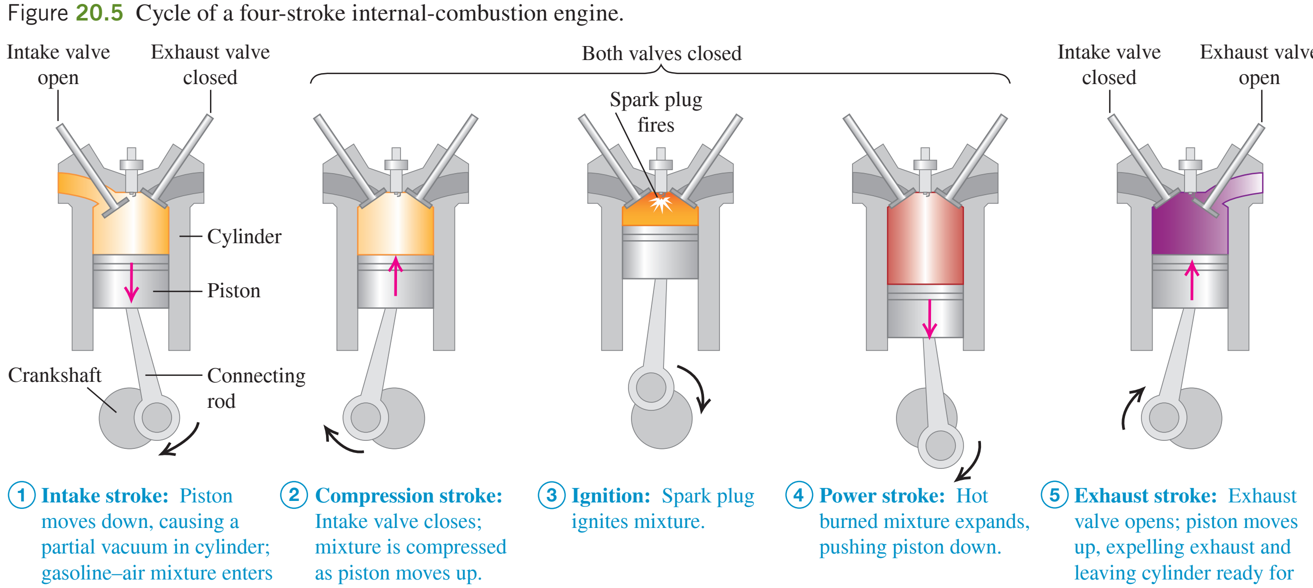
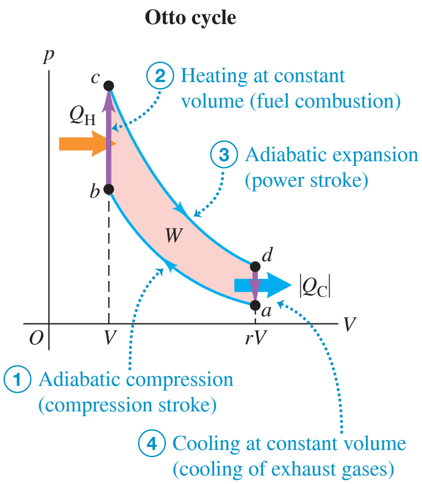


Figure **20.6** The pV-diagram for the Otto cycle, an idealized model of the thermodynamic processes in a gasoline engine.



We can calculate the efficiency of this idealized cycle. Processes bc and da are constant-volume, so the heats Q_H and Q_C are related simply to the temperatures at points a, b, c, and d:

$$Q_H = nC_V(T_c - T_b) > 0$$

$$Q_C = nC_V(T_a - T_d) < 0$$

The thermal efficiency is given by Eq. (20.4). Inserting the above expressions and canceling out the common factor nC_V, we find

$$e = \frac{Q_H + Q_C}{Q_H} = \frac{T_c - T_b + T_a - T_d}{T_c - T_b} \tag{20.5}$$

To simplify this further, we use the temperature–volume relationship for adiabatic processes for an ideal gas, Eq. (19.22). For the two adiabatic processes ab and cd,

$$T_a(rV)^{\gamma-1} = T_b V^{\gamma-1} \quad \text{and} \quad T_d(rV)^{\gamma-1} = T_c V^{\gamma-1}$$

where γ is the ratio of heat capacities for the gas in the engine (see Section 19.7). We divide each of these equations by the common factor $V^{\gamma-1}$ and substitute the resulting expressions for T_b and T_c back into Eq. (20.5). The result is

$$e = \frac{T_d r^{\gamma-1} - T_a r^{\gamma-1} + T_a - T_d}{T_d r^{\gamma-1} - T_a r^{\gamma-1}} = \frac{(T_d - T_a)(r^{\gamma-1} - 1)}{(T_d - T_a) r^{\gamma-1}}$$

Dividing out the common factor $(T_d - T_a)$, we get

Thermal efficiency in Otto cycle

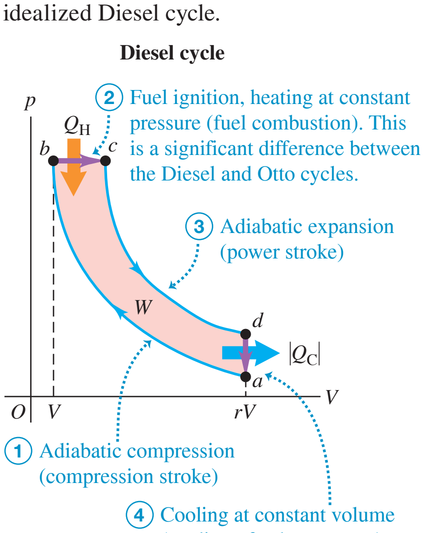
$$e = 1 - \frac{1}{r^{\gamma-1}} \tag{20.6}$$

Compression ratio (r); Ratio of heat capacities (γ)

The thermal efficiency given by Eq. (20.6) is always less than unity, even for this idealized model. With $r = 8$ and $\gamma = 1.4$ (the value for air) the theoretical efficiency is $e = 0.56$, or 56%. The efficiency can be increased by increasing r. However, this also increases the temperature at the end of the adiabatic compression of the air–fuel mixture. If the temperature is too high, the mixture explodes spontaneously during compression instead of burning evenly after the spark plug ignites it. This is called *pre-ignition* or *detonation;* it causes a knocking sound and can damage the engine. The octane rating of a gasoline is a measure of its antiknock qualities. The maximum practical compression ratio for high-octane, or "premium," gasoline is about 10 to 13.

The Otto cycle is a highly idealized model. It assumes that the mixture behaves as an ideal gas; it ignores friction, turbulence, loss of heat to cylinder walls, and many other effects that reduce the efficiency of an engine. Efficiencies of real gasoline engines are typically around 35%.

Figure **20.7** The pV-diagram for the idealized Diesel cycle.



The Diesel Cycle

The Diesel engine is similar in operation to the gasoline engine. The most important difference is that there is no fuel in the cylinder at the beginning of the compression stroke. A little before the beginning of the power stroke, the injectors start to inject fuel directly into the cylinder, just fast enough to keep the pressure approximately constant during the first part of the power stroke. Because of the high temperature developed during the adiabatic compression, the fuel ignites spontaneously as it is injected; no spark plugs are needed.

Figure 20.7 shows the idealized **Diesel cycle.** Starting at point a, air is compressed adiabatically to point b, heated at constant pressure to point c, expanded adiabatically to point d, and cooled at constant volume to point a. Because there is no fuel in the cylinder during most of the compression stroke, pre-ignition cannot occur, and the compression ratio r can be much higher than for a gasoline engine. This improves efficiency and ensures reliable ignition when the fuel is injected (because of the high temperature reached during the adiabatic compression). Values of r of 15 to 20 are typical; with these values and $\gamma = 1.4$, the theoretical efficiency of the idealized Diesel cycle is about 0.65 to 0.70. As with the Otto cycle, the efficiency of any actual engine is substantially less than this. While Diesel engines are very efficient, they must be built to much tighter tolerances than gasoline engines and the fuel-injection system requires careful maintenance.

TEST YOUR UNDERSTANDING OF SECTION 20.3 For an Otto-cycle engine with cylinders of a fixed size and a fixed compression ratio, which of the following aspects of the pV-diagram in Fig. 20.6 would change if you doubled the amount of fuel burned per cycle? (There may be more than one correct answer.) (i) The vertical distance between points b and c; (ii) the vertical distance between points a and d; (iii) the horizontal distance between points b and a.

ANSWER

❙ **(i), (ii)** Doubling the amount of fuel burned per cycle means that Q_H is doubled, so the resulting pressure increase from b to c in Fig. 20.6 is greater. The compression ratio and hence the efficiency remain the same, so $|Q_C|$ (the amount of heat rejected to the environment) must increase by the same factor as Q_H. Hence the pressure drop from d to a in Fig. 20.6 is also greater. The volume V and the compression ratio r don't change, so the horizontal dimensions of the pV-diagram don't change.

20.4 REFRIGERATORS

We can think of a **refrigerator** as a heat engine operating in reverse. A heat engine takes heat from a hot place and gives off heat to a colder place. A refrigerator does the opposite; it takes heat from a cold place (the inside of the refrigerator) and gives it off to a warmer place (usually the air in the room where the refrigerator is located). A heat engine has a net *output* of mechanical work; the refrigerator requires a net *input* of mechanical work. With the sign conventions from Section 20.2, for a refrigerator Q_C is positive but both W and Q_H are negative; hence $|W| = -W$ and $|Q_H| = -Q_H$.

Figure 20.8 shows an energy-flow diagram for a refrigerator. From the first law for a cyclic process,

$$Q_H + Q_C - W = 0 \qquad \text{or} \qquad -Q_H = Q_C - W$$

or, because both Q_H and W are negative,

$$|Q_H| = Q_C + |W| \tag{20.7}$$

Thus, as the diagram shows, the heat $|Q_H|$ leaving the working substance and given to the hot reservoir is always *greater* than the heat Q_C taken from the cold reservoir. Note that the absolute-value relationship

$$|Q_H| = |Q_C| + |W| \tag{20.8}$$

is valid for both heat engines and refrigerators.

From an economic point of view, the best refrigeration cycle is one that removes the greatest amount of heat $|Q_C|$ from the inside of the refrigerator for the least expenditure of mechanical work, $|W|$. The relevant ratio is therefore $|Q_C|/|W|$; the larger this ratio, the better the refrigerator. We call this ratio the **coefficient of performance,** K. From Eq. (20.8), $|W| = |Q_H| - |Q_C|$, so

Coefficient of performance of a refrigerator

$$K = \frac{|Q_C|}{|W|} = \frac{|Q_C|}{|Q_H| - |Q_C|} \tag{20.9}$$

Heat removed from inside of refrigerator

Work input of refrigerator

Heat rejected to outside air

As always, we measure Q_H, Q_C, and W all in the same energy units; K is then a dimensionless number.

Figure **20.8** Schematic energy-flow diagram of a refrigerator.

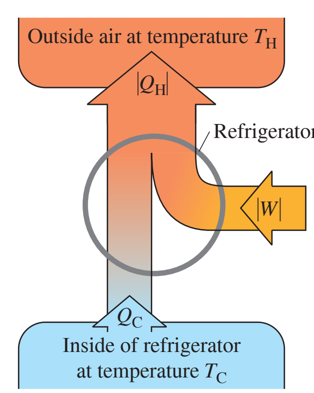
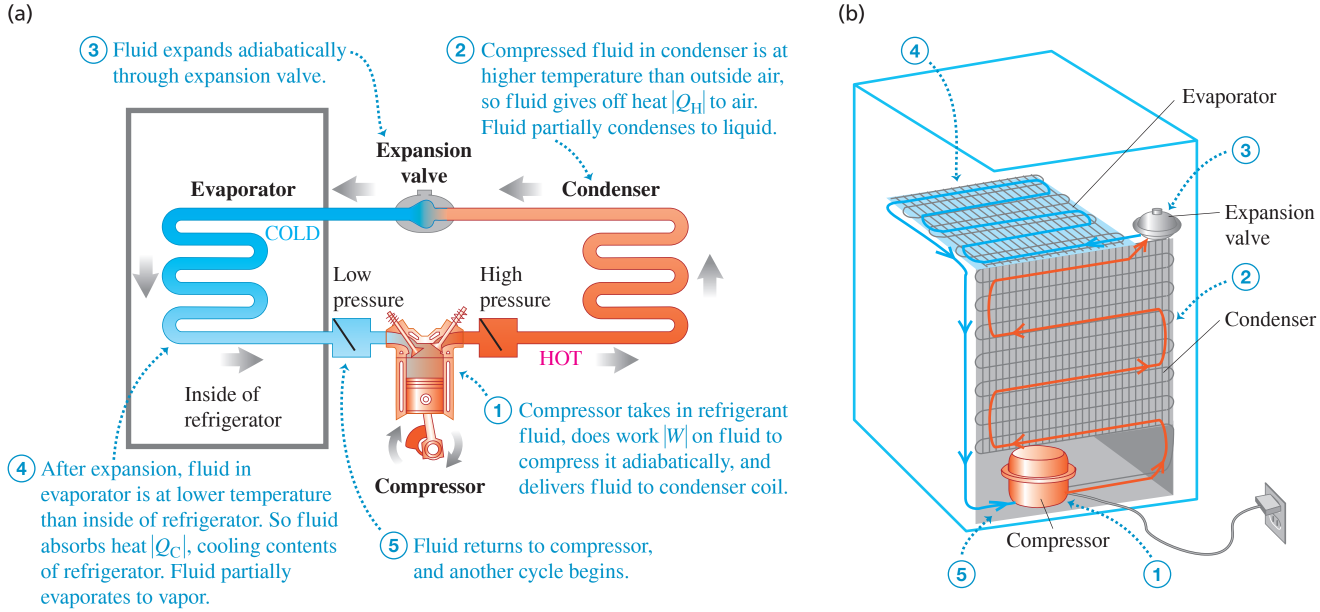


Figure **20.9** (a) Principle of the mechanical refrigeration cycle. (b) How the key elements are arranged in a practical refrigerator.

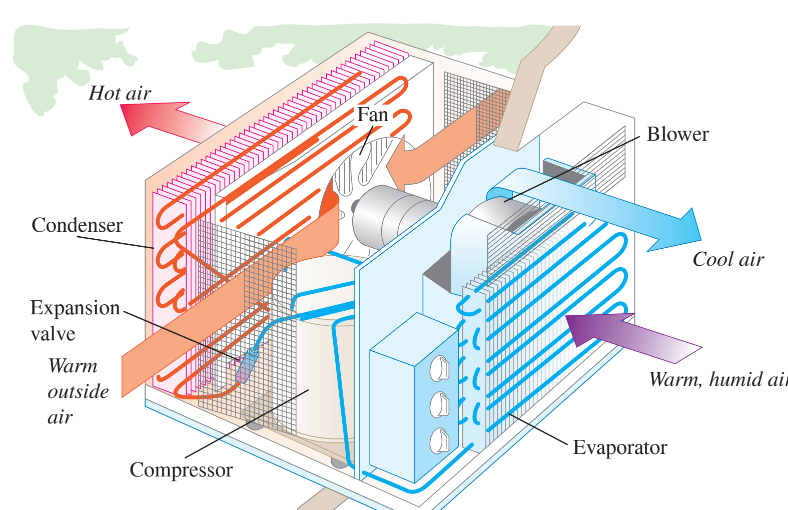


Practical Refrigerators

Figure 20.9a shows the principles of the cycle used in home refrigerators. The fluid “circuit” contains a refrigerant fluid (the working substance). The left side of the circuit (including the cooling coils inside the refrigerator) is at low temperature and low pressure; the right side (including the condenser coils outside the refrigerator) is at high temperature and high pressure. Ordinarily, both sides contain liquid and vapor in phase equilibrium. In each cycle the fluid absorbs heat $|Q_C|$ from the inside of the refrigerator on the left side and gives off heat $|Q_H|$ to the surrounding air on the right side. The compressor, usually driven by an electric motor (Fig. 20.9b), does work $|W|$ *on* the fluid during each cycle. So the compressor requires energy input, which is why refrigerators have to be plugged in.

An air conditioner operates on exactly the same principle. In this case the refrigerator box becomes a room or an entire building. The evaporator coils are inside, the condenser is outside, and fans circulate air through these (**Fig. 20.10**). In large installations the condenser coils are often cooled by water. For air conditioners the quantities of greatest practical importance are the *rate* of heat removal (the heat current H from the region being

Figure **20.10** An air conditioner works on the same principle as a refrigerator.



cooled) and the *power* input $P = W/t$ to the compressor. If heat $|Q_C|$ is removed in time t, then $H = |Q_C|/t$. Then we can express the coefficient of performance as

$$K = \frac{|Q_C|}{|W|} = \frac{Ht}{Pt} = \frac{H}{P}$$

Typical room air conditioners have heat removal rates H of about 1500–3000 W and require electric power input of about 600 to 1200 W. Typical coefficients of performance are about 3; the actual values depend on the inside and outside temperatures.

A variation on this theme is the **heat pump,** used to heat buildings by cooling the outside air. It functions like a refrigerator turned inside out. The evaporator coils are outside, where they take heat from cold air, and the condenser coils are inside, where they give off heat to the warmer air. With proper design, the heat $|Q_H|$ delivered to the inside per cycle can be considerably greater than the work $|W|$ required to get it there.

Work is *always* needed to transfer heat from a colder to a hotter object. Heat flows spontaneously from hotter to colder, and to reverse this flow requires the addition of work from the outside. Experience shows that it is impossible to make a refrigerator that transports heat from a colder object to a hotter object without the addition of work. If no work were needed, the coefficient of performance would be infinite. We call such a device a *workless refrigerator;* it is a mythical beast, like the unicorn and the free lunch. ?

TEST YOUR UNDERSTANDING OF SECTION 20.4 Can you cool your house by leaving the refrigerator door open?

ANSWER

no A refrigerator uses an input of work to transfer heat from one system (the refrigerator's interior) to another system (its exterior, which includes the house in which the refrigerator is installed). If the door is open, these two systems are really the *same* system and will eventually come to the same temperature. By the first law of thermodynamics, all of the work input to the refrigerator motor will be converted into heat and the temperature in your house will actually *increase*. To cool the house you need a system that will transfer heat from it to the outside world, such as an air conditioner or heat pump.

20.5 THE SECOND LAW OF THERMODYNAMICS

Experimental evidence suggests strongly that it is *impossible* to build a heat engine that converts heat completely to work—that is, an engine with 100% thermal efficiency. This impossibility is the basis of one statement of the **second law of thermodynamics,** as follows:

SECOND LAW OF THERMODYNAMICS ("engine" statement) It is impossible for any system to undergo a process in which it absorbs heat from a reservoir at a single temperature and converts the heat completely into mechanical work, with the system ending in the same state in which it began.

This is also known to physicists as the *Kelvin–Planck statement* of the second law.

The basis of the second law of thermodynamics is the difference between the nature of internal energy and that of macroscopic mechanical energy. In a moving object the molecules have random motion, but superimposed on this is a coordinated motion of every molecule in the direction of the object's velocity. The kinetic energy associated with this *coordinated* macroscopic motion is what we call the kinetic energy of the moving object. The kinetic and potential energies associated with the *random* motion constitute the internal energy.

When an object sliding on a surface comes to rest as a result of friction, the organized motion of the object is converted to random motion of molecules in the object and in the surface. Since we cannot control the motions of individual molecules, we cannot convert this random motion completely back to organized motion. We can convert *part* of it, and this is what a heat engine does.

If the second law were *not* true, we could power an automobile or run a power plant by cooling the surrounding air. Neither of these impossibilities violates the *first* law of thermodynamics. The second law, therefore, is not a deduction from the first but stands by itself as a separate law of nature. The first law denies the possibility of creating or destroying energy; the second law limits the *availability* of energy and the ways in which it can be used and converted.

Restating the Second Law

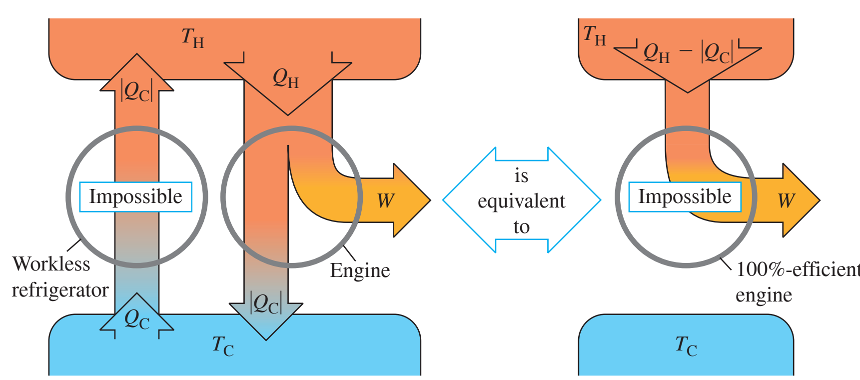
Our analysis of refrigerators in Section 20.4 forms the basis for an alternative statement of the second law of thermodynamics. Heat flows spontaneously from hotter to colder objects, never the reverse. A refrigerator does take heat from a colder to a hotter object, but its operation requires an input of mechanical energy or work. We can generalize this observation:

SECOND LAW OF THERMODYNAMICS (“refrigerator” statement) It is impossible for any process to have as its sole result the transfer of heat from a cooler to a hotter object.

This is also known as the *Clausius statement* of the second law. It may not seem to be very closely related to the “engine” statement. In fact, though, the two statements are completely equivalent. For example, if we could build a workless refrigerator, violating the second or “refrigerator” statement of the second law, we could use it in conjunction with a heat engine, pumping the heat rejected by the engine back to the hot reservoir to be reused. This composite machine (**Fig. 20.11a**) would violate the “engine” statement of the second

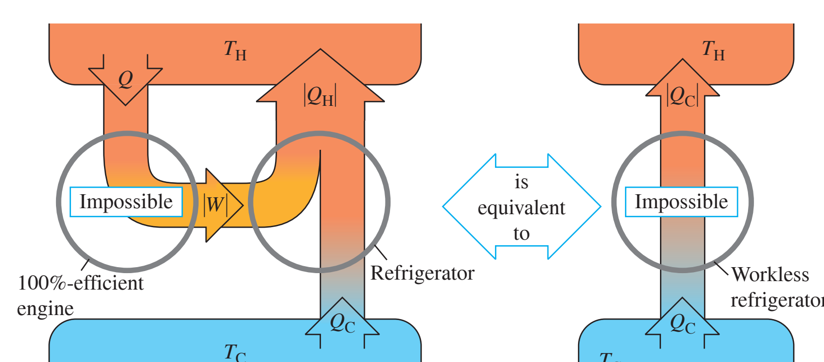
Figure 20.11 Energy-flow diagrams showing that the two forms of the second law are equivalent.

(a) The “engine” statement of the second law of thermodynamics



If a workless refrigerator were possible, it could be used in conjunction with an ordinary heat engine to form a 100%-efficient engine, converting heat $Q_H - |Q_C|$ completely to work.

(b) The “refrigerator” statement of the second law of thermodynamics



If a 100%-efficient engine were possible, it could be used in conjunction with an ordinary refrigerator to form a workless refrigerator, transferring heat Q_C from the cold to the hot reservoir with no input of work.

law because its net effect would be to take a net quantity of heat $Q_H - |Q_C|$ from the hot reservoir and convert it completely to work W.

Alternatively, if we could make an engine with 100% thermal efficiency, in violation of the first statement, we could run it using heat from the hot reservoir and use the work output to drive a refrigerator that pumps heat from the cold reservoir to the hot (Fig. 20.11b). This composite device would violate the "refrigerator" statement because its net effect would be to take heat Q_C from the cold reservoir and deliver it to the hot reservoir without requiring any input of work. Thus any device that violates one form of the second law can be used to make a device that violates the other form. If violations of the first form are impossible, so are violations of the second!

The conversion of work to heat and the heat flow from hot to cold across a finite temperature gradient are *irreversible* processes. The "engine" and "refrigerator" statements of the second law state that these processes can be only partially reversed. We could cite other examples. Gases naturally flow from a region of high pressure to a region of low pressure; gases and miscible liquids left by themselves always tend to mix, not to unmix. The second law of thermodynamics is an expression of the inherent one-way aspect of these and many other irreversible processes. Energy conversion is an essential aspect of all plant and animal life and of human technology, so the second law of thermodynamics is of fundamental importance.

TEST YOUR UNDERSTANDING OF SECTION 20.5 Would a 100%-efficient engine (Fig. 20.11a) violate the *first* law of thermodynamics? What about a workless refrigerator (Fig. 20.11b)?

ANSWER

no, no Both the 100%-efficient engine of Fig. 20.11a and the workless refrigerator of Fig. 20.11b return to the same state at the end of a cycle as at the beginning, so the net change in internal energy of each system is zero $(\Delta U = 0)$. For the 100%-efficient engine, the net heat flow into the engine equals the net work done, so $Q = W$, $Q - W = 0$, and the first law $(\Delta U = Q - W)$ is obeyed. For the workless refrigerator, no net work is done (so $W = 0$) and as much heat flows into it as out (so $Q = 0$), so again $Q - W = 0$ and $\Delta U = Q - W$ in accordance with the first law. It is the *second* law of thermodynamics that tells us that both the 100%-efficient engine and the workless refrigerator are impossible.

20.6 THE CARNOT CYCLE

According to the second law, no heat engine can have 100% efficiency. How great an efficiency *can* an engine have, given two heat reservoirs at temperatures T_H and T_C? This question was answered in 1824 by the French engineer Sadi Carnot (1796–1832), who developed a hypothetical, idealized heat engine that has the maximum possible efficiency consistent with the second law. The cycle of this engine is called the **Carnot cycle.**

To understand the rationale of the Carnot cycle, we return to *reversibility* and its relationship to directions of thermodynamic processes. Conversion of work to heat is an irreversible process; the purpose of a heat engine is a *partial* reversal of this process, the conversion of heat to work with as great an efficiency as possible. For maximum heat-engine efficiency, therefore, *we must avoid all irreversible processes* (**Fig. 20.12**).

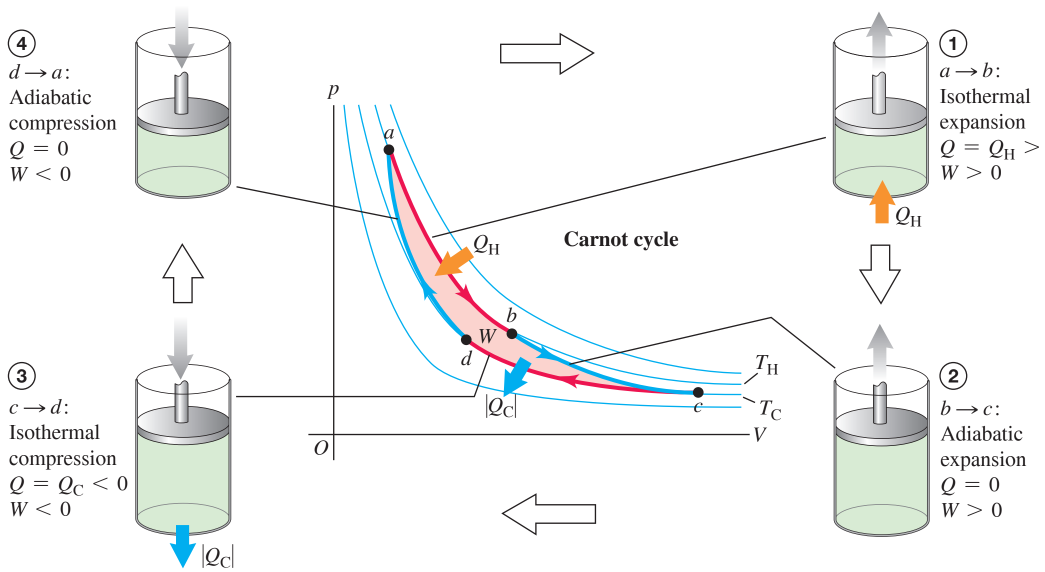
Heat flow through a finite temperature drop is an irreversible process. Therefore, during heat transfer in the Carnot cycle there must be *no* finite temperature difference. When the engine takes heat from the hot reservoir at temperature T_H, the working substance of the engine must also be at T_H; otherwise, irreversible heat flow would occur. Similarly, when the engine discards heat to the cold reservoir at T_C, the engine itself must be at T_C. That is, every process that involves heat transfer must be *isothermal* at either T_H or T_C.

Conversely, in any process in which the temperature of the working substance of the engine is intermediate between T_H and T_C, there must be *no* heat transfer between the engine and either reservoir because such heat transfer could not be reversible. Therefore any process in which the temperature T of the working substance changes must be *adiabatic*.

The bottom line is that every process in our idealized cycle must be either isothermal or adiabatic. In addition, thermal and mechanical equilibrium must be maintained at all times so that each process is completely reversible.

Figure 20.12 The temperature of the firebox of a steam engine is much higher than the temperature of water in the boiler, so heat flows irreversibly from firebox to water. Carnot's quest to understand the efficiency of steam engines led him to the idea that an ideal engine would involve only *reversible* processes.

Figure **20.13** The Carnot cycle for an ideal gas. The light blue lines in the pV-diagram are isotherms (curves of constant temperature) and the dark blue lines are adiabats (curves of zero heat flow).



Steps of the Carnot Cycle

The Carnot cycle consists of two reversible isothermal and two reversible adiabatic processes. **Figure 20.13** shows a Carnot cycle using as its working substance an ideal gas in a cylinder with a piston. It consists of the following steps:

1. The gas expands isothermally at temperature T_H, absorbing heat Q_H (ab).
2. It expands adiabatically until its temperature drops to T_C (bc).
3. It is compressed isothermally at T_C, rejecting heat $|Q_C|$ (cd).
4. It is compressed adiabatically back to its initial state at temperature T_H (da).

We can calculate the thermal efficiency e of a Carnot engine in the special case shown in Fig. 20.13 in which the working substance is an *ideal gas*. To do this, we'll first find the ratio Q_C/Q_H of the quantities of heat transferred in the two isothermal processes and then use Eq. (20.4) to find e.

For an ideal gas the internal energy U depends only on temperature and is thus constant in any isothermal process. For the isothermal expansion ab, $\Delta U_{ab} = 0$ and Q_H is equal to the work W_{ab} done by the gas during its isothermal expansion at temperature T_H. We calculated this work in Example 19.1 (Section 19.2); using that result, we have

$$Q_H = W_{ab} = nRT_H \ln \frac{V_b}{V_a} \tag{20.10}$$

Similarly,

$$Q_C = W_{cd} = nRT_C \ln \frac{V_d}{V_c} = -nRT_C \ln \frac{V_c}{V_d} \tag{20.11}$$

Because V_d is less than V_c, Q_C is negative ($Q_C = -|Q_C|$); heat flows out of the gas during the isothermal compression at temperature T_C.

The ratio of the two quantities of heat is thus

$$\frac{Q_C}{Q_H} = -\left(\frac{T_C}{T_H}\right) \frac{\ln (V_c/V_d)}{\ln (V_b/V_a)} \tag{20.12}$$

This can be simplified further by use of the temperature–volume relationship for an adiabatic process, Eq. (19.22). We find for the two adiabatic processes:

$$T_H V_b^{\gamma-1} = T_C V_c^{\gamma-1} \qquad \text{and} \qquad T_H V_a^{\gamma-1} = T_C V_d^{\gamma-1}$$

Dividing the first of these by the second, we find

$$\frac{V_b^{\gamma-1}}{V_a^{\gamma-1}} = \frac{V_c^{\gamma-1}}{V_d^{\gamma-1}} \qquad \text{and} \qquad \frac{V_b}{V_a} = \frac{V_c}{V_d}$$

Thus the two logarithms in Eq. (20.12) are equal, and that equation reduces to

$$\frac{Q_C}{Q_H} = -\frac{T_C}{T_H} \quad \text{or} \quad \frac{|Q_C|}{|Q_H|} = \frac{T_C}{T_H} \qquad \text{(heat transfer in a Carnot engine)} \tag{20.13}$$

The ratio of the heat rejected at T_C to the heat absorbed at T_H is just equal to the ratio T_C/T_H. Then from Eq. (20.4) the efficiency of the Carnot engine is

Efficiency of a Carnot engine

Temperature of cold reservoir

$$e_{\text{Carnot}} = 1 - \frac{T_C}{T_H} = \frac{T_H - T_C}{T_H} \tag{20.14}$$

Temperature of hot reservoir

This simple result says that the efficiency of a Carnot engine depends only on the temperatures of the two heat reservoirs. The efficiency is large when the temperature *difference* is large, and it is very small when the temperatures are nearly equal. The efficiency can never be exactly unity unless $T_C = 0$; we'll see later that this, too, is impossible.

CAUTION Use Kelvin temperature in all Carnot calculations In Carnot-cycle calculations, you must use *absolute* (Kelvin) temperatures only. That's because Eqs. (20.10) through (20.14) come from the ideal-gas equation $pV = nRT$, in which T is absolute temperature.

EXAMPLE 20.2 Analyzing a Carnot engine I

WITH VARIATION PROBLEMS

A Carnot engine takes 2000 J of heat from a reservoir at 500 K, does some work, and discards some heat to a reservoir at 350 K. How much work does it do, how much heat is discarded, and what is its efficiency?

IDENTIFY and SET UP This problem involves a Carnot engine, so we can use the ideas of this section and those of Section 20.2 (which apply to heat engines of all kinds). **Figure 20.14** shows the energy-flow diagram. We have $Q_H = 2000\ \text{J}$, $T_H = 500\ \text{K}$, and $T_C = 350\ \text{K}$. We use Eq. (20.13) to find Q_C, and then use the first law of thermodynamics as given by Eq. (20.2) to find W. We find the efficiency e from T_C and T_H from Eq. (20.14).

EXECUTE From Eq. (20.13),

$$Q_C = -Q_H \frac{T_C}{T_H} = -(2000\ \text{J})\frac{350\ \text{K}}{500\ \text{K}} = -1400\ \text{J}$$

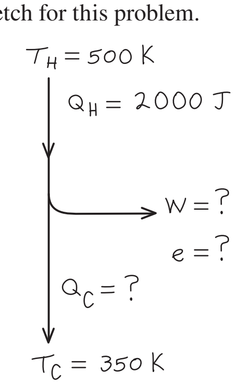
Then from Eq. (20.2), the work done is

$$W = Q_H + Q_C = 2000\ \text{J} + (-1400\ \text{J}) = 600\ \text{J}$$

From Eq. (20.14), the thermal efficiency is

$$e = 1 - \frac{T_C}{T_H} = 1 - \frac{350\ \text{K}}{500\ \text{K}} = 0.30 = 30\%$$

Figure 20.14 Our sketch for this problem.



EVALUATE The negative sign of Q_C is correct: It shows that 1400 J of heat flows *out* of the engine and into the cold reservoir. We can check our result for e by using the basic definition of thermal efficiency, Eq. (20.3):

$$e = \frac{W}{Q_H} = \frac{600\ \text{J}}{2000\ \text{J}} = 0.30 = 30\%$$

KEYCONCEPT An engine that utilizes the Carnot cycle, which consists of only reversible isothermal and reversible adiabatic processes, has the greatest efficiency possible consistent with the second law of thermodynamics. The smaller the ratio of the temperature of the cold reservoir to the temperature of the hot reservoir, the greater the Carnot efficiency.

EXAMPLE 20.3 Analyzing a Carnot engine II

WITH VARIATION PROBLEMS

Suppose 0.200 mol of an ideal diatomic gas ($\gamma = 1.40$) undergoes a Carnot cycle between 227°C and 27°C, starting at $p_a = 10.0 \times 10^5$ Pa at point a in the pV-diagram of Fig. 20.13. The volume doubles during the isothermal expansion step $a \rightarrow b$. (a) Find the pressure and volume at points a, b, c, and d. (b) Find Q, W, and ΔU for each step and for the entire cycle. (c) Find the efficiency directly from the results of part (b), and compare with the value calculated from Eq. (20.14).

IDENTIFY and SET UP This problem involves the properties of the Carnot cycle and those of an ideal gas. We are given the number of moles n and the pressure and temperature at point a (which is at the higher of the two reservoir temperatures); we can find the volume at a from the ideal-gas equation $pV = nRT$. We then find the pressure and volume at points b, c, and d from the known doubling of volume in step $a \rightarrow b$, from equations given in this section, and from $pV = nRT$. In each step we use Eqs. (20.10) and (20.11) to find the heat flow and work done and Eq. (19.13) to find the internal energy change.

EXECUTE (a) With $T_H = (227 + 273.15)$ K $= 500$ K and $T_C = (27 + 273.15)$ K $= 300$ K, $pV = nRT$ yields

$$V_a = \frac{nRT_H}{p_a} = \frac{(0.200 \text{ mol})(8.314 \text{ J/mol}\cdot\text{K})(500 \text{ K})}{10.0 \times 10^5 \text{ Pa}}$$
$$= 8.31 \times 10^{-4} \text{ m}^3$$

The volume doubles during the isothermal expansion $a \rightarrow b$:

$$V_b = 2V_a = 2(8.31 \times 10^{-4} \text{ m}^3) = 16.6 \times 10^{-4} \text{ m}^3$$

Because the expansion $a \rightarrow b$ is isothermal, $p_aV_a = p_bV_b$, so

$$p_b = \frac{p_aV_a}{V_b} = 5.00 \times 10^5 \text{ Pa}$$

For the adiabatic expansion $b \rightarrow c$, we use the equation $T_H V_b^{\gamma-1} = T_C V_c^{\gamma-1}$ that follows Eq. (20.12) as well as the ideal-gas equation:

$$V_c = V_b\left(\frac{T_H}{T_C}\right)^{1/(\gamma-1)} = (16.6 \times 10^{-4} \text{ m}^3)\left(\frac{500 \text{ K}}{300 \text{ K}}\right)^{2.5}$$
$$= 59.6 \times 10^{-4} \text{ m}^3$$
$$p_c = \frac{nRT_C}{V_c} = \frac{(0.200 \text{ mol})(8.314 \text{ J/mol}\cdot\text{K})(300 \text{ K})}{59.6 \times 10^{-4} \text{ m}^3}$$
$$= 0.837 \times 10^5 \text{ Pa}$$

For the adiabatic compression $d \rightarrow a$, we have $T_C V_d^{\gamma-1} = T_H V_a^{\gamma-1}$ and so

$$V_d = V_a\left(\frac{T_H}{T_C}\right)^{1/(\gamma-1)} = (8.31 \times 10^{-4} \text{ m}^3)\left(\frac{500 \text{ K}}{300 \text{ K}}\right)^{2.5}$$
$$= 29.8 \times 10^{-4} \text{ m}^3$$
$$p_d = \frac{nRT_C}{V_d} = \frac{(0.200 \text{ mol})(8.314 \text{ J/mol}\cdot\text{K})(300 \text{ K})}{29.8 \times 10^{-4} \text{ m}^3}$$
$$= 1.67 \times 10^5 \text{ Pa}$$

(b) For the isothermal expansion $a \rightarrow b$, $\Delta U_{ab} = 0$. From Eq. (20.10),

$$W_{ab} = Q_H = nRT_H \ln\frac{V_b}{V_a}$$
$$= (0.200 \text{ mol})(8.314 \text{ J/mol}\cdot\text{K})(500 \text{ K})(\ln 2) = 576 \text{ J}$$

For the adiabatic expansion $b \rightarrow c$, $Q_{bc} = 0$. From the first law of thermodynamics, $\Delta U_{bc} = Q_{bc} - W_{bc} = -W_{bc}$; the work W_{bc} done by the gas in this adiabatic expansion equals the negative of the change in internal energy of the gas. From Eq. (19.13) we have $\Delta U = nC_V\Delta T$, where $\Delta T = T_C - T_H$. Using $C_V = 20.8$ J/mol·K for an ideal diatomic gas, we find

$$W_{bc} = -\Delta U_{bc} = -nC_V(T_C - T_H) = nC_V(T_H - T_C)$$
$$= (0.200 \text{ mol})(20.8 \text{ J/mol}\cdot\text{K})(500 \text{ K} - 300 \text{ K}) = 832 \text{ J}$$

For the isothermal compression $c \rightarrow d$, $\Delta U_{cd} = 0$; Eq. (20.11) gives

$$W_{cd} = Q_C = nRT_C \ln\frac{V_d}{V_c}$$
$$= (0.200 \text{ mol})(8.314 \text{ J/mol}\cdot\text{K})(300 \text{ K})\left(\ln\frac{29.8 \times 10^{-4} \text{ m}^3}{59.6 \times 10^{-4} \text{ m}^3}\right)$$
$$= -346 \text{ J}$$

For the adiabatic compression $d \rightarrow a$, $Q_{da} = 0$ and

$$W_{da} = -\Delta U_{da} = -nC_V(T_H - T_C) = nC_V(T_C - T_H)$$
$$= (0.200 \text{ mol})(20.8 \text{ J/mol}\cdot\text{K})(300 \text{ K} - 500 \text{ K}) = -832 \text{ J}$$

We can tabulate these results as follows:

Process	Q	W	ΔU
$a \rightarrow b$	576 J	576 J	0
$b \rightarrow c$	0	832 J	−832 J
$c \rightarrow d$	−346 J	−346 J	0
$d \rightarrow a$	0	−832 J	832 J
Total	230 J	230 J	0

(c) From the above table, $Q_H = 576$ J and the total work is 230 J. Thus

$$e = \frac{W}{Q_H} = \frac{230 \text{ J}}{576 \text{ J}} = 0.40 = 40\%$$

We can compare this to the result from Eq. (20.14),

$$e = \frac{T_H - T_C}{T_H} = \frac{500 \text{ K} - 300 \text{ K}}{500 \text{ K}} = 0.40 = 40\%$$

EVALUATE The table in part (b) shows that for the entire cycle $Q = W$ and $\Delta U = 0$, just as we would expect: In a complete cycle, the *net* heat input is used to do work, and there is zero net change in the internal energy of the system. Note also that the quantities of work in the two adiabatic processes are negatives of each other. Can you show from the analysis leading to Eq. (20.13) that this must *always* be the case in a Carnot cycle?

KEYCONCEPT When analyzing any of the steps taken by a heat engine as it goes through a thermodynamic cycle, if you can calculate any two of the quantities ΔU, Q, and W, you can calculate the third using the first law of thermodynamics, $\Delta U = Q - W$. The efficiency of the engine equals the sum of the work done in all steps divided by the heat taken in from the hot reservoir.

The Carnot Refrigerator

Because each step in the Carnot cycle is reversible, the *entire cycle* may be reversed, converting the engine into a refrigerator. The coefficient of performance of the Carnot refrigerator is obtained by combining the general definition of K, Eq. (20.9), with Eq. (20.13) for the Carnot cycle. We first rewrite Eq. (20.9) as

$$K = \frac{|Q_C|}{|Q_H| - |Q_C|}$$
$$= \frac{|Q_C|/|Q_H|}{1 - |Q_C|/|Q_H|}$$

Then we substitute Eq. (20.13), $|Q_C|/|Q_H| = T_C/T_H$, into this expression:

Coefficient of performance of a Carnot refrigerator

$$K_{\text{Carnot}} = \frac{T_C}{T_H - T_C} \qquad (20.15)$$

Temperature of cold reservoir

Temperature of hot reservoir

When the temperature difference $T_H - T_C$ is small, K is much larger than unity; in this case a lot of heat can be "pumped" from the lower to the higher temperature with only a little expenditure of work. But the greater the temperature difference, the smaller the value of K and the more work is required to transfer a given quantity of heat.

EXAMPLE 20.4 Analyzing a Carnot refrigerator

WITH VARIATION PROBLEMS

If the cycle described in Example 20.3 is run backward as a refrigerator, what is its coefficient of performance?

IDENTIFY and SET UP This problem uses the ideas of Section 20.3 (for refrigerators in general) and the above discussion of Carnot refrigerators. Equation (20.9) gives the coefficient of performance K of *any* refrigerator in terms of the heat Q_C extracted from the cold reservoir per cycle and the work W that must be done per cycle.

EXECUTE In Example 20.3 we found that in one cycle the Carnot engine rejects heat $Q_C = -346$ J to the cold reservoir and does work $W = 230$ J. When run in reverse as a refrigerator, the system extracts heat $Q_C = +346$ J from the cold reservoir while requiring a work input of $W = -230$ J. From Eq. (20.9),

$$K = \frac{|Q_C|}{|W|} = \frac{346 \text{ J}}{230 \text{ J}}$$
$$= 1.50$$

Because this is a Carnot cycle, we can also use Eq. (20.15):

$$K = \frac{T_C}{T_H - T_C} = \frac{300 \text{ K}}{500 \text{ K} - 300 \text{ K}}$$
$$= 1.50$$

EVALUATE Equations (20.14) and (20.15) show that e and K for a Carnot cycle depend only on T_H and T_C, and we don't need to calculate Q and W. For cycles containing irreversible processes, however, these two equations are not valid, and more detailed calculations are necessary.

KEYCONCEPT A refrigerator takes heat in from a cold reservoir and rejects heat to a hot reservoir. Work must be done on the refrigerator to make this happen. The coefficient of performance K of the refrigerator equals the amount of heat that is rejected to the hot reservoir divided by the amount of work that must be done on the refrigerator. The greater the value of K, the more heat can be removed from the cold reservoir for a given expenditure of work.

The Carnot Cycle and the Second Law

We can prove that **no engine can be more efficient than a Carnot engine operating between the same two temperatures.** The key to the proof is the above observation that since each step in the Carnot cycle is reversible, the *entire cycle* may be reversed. Run backward, the engine becomes a refrigerator. Suppose we have an engine that is more

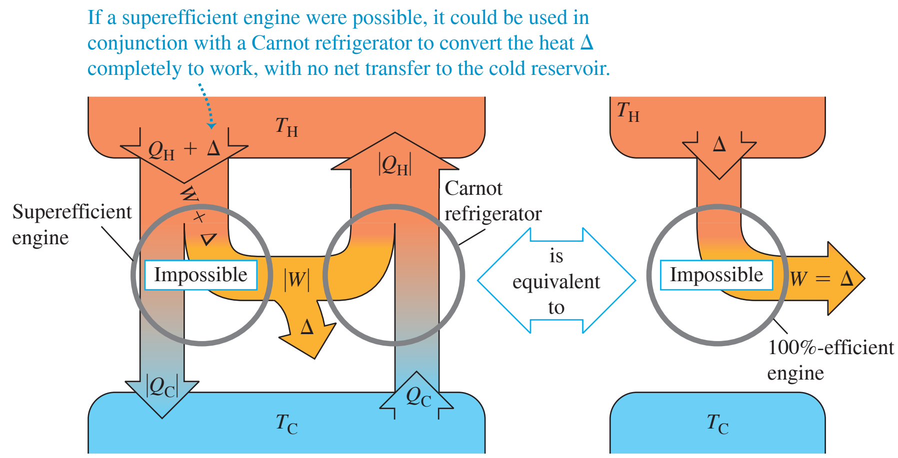


Figure **20.15** Proving that the Carnot engine has the highest possible efficiency. A "superefficient" engine (more efficient than a Carnot engine) combined with a Carnot refrigerator could convert heat completely into work with no net heat transfer to the cold reservoir. This would violate the second law of thermodynamics.

efficient than a Carnot engine (**Fig. 20.15**). Let the Carnot engine, run backward as a refrigerator by negative work $-|W|$, take in heat Q_C from the cold reservoir and expel heat $|Q_H|$ to the hot reservoir. The superefficient engine expels heat $|Q_C|$, but to do this, it takes in a greater amount of heat $Q_H + \Delta$. Its work output is then $W + \Delta$, and the net effect of the two machines together is to take a quantity of heat Δ and convert it completely into work. This violates the "engine" statement of the second law. We could construct a similar argument that a superefficient engine could be used to violate the "refrigerator" statement of the second law. Note that we don't have to assume that the superefficient engine is reversible. In a similar way we can show that *no refrigerator can have a greater coefficient of performance than a Carnot refrigerator operating between the same two temperatures.*

Thus the statement that no engine can be more efficient than a Carnot engine is yet another equivalent statement of the second law of thermodynamics. It also follows directly that **all Carnot engines operating between the same two temperatures have the same efficiency, irrespective of the nature of the working substance.** Although we derived Eq. (20.14) for a Carnot engine by using an ideal gas as its working substance, it is in fact valid for *any* Carnot engine, no matter what its working substance.

Equation (20.14), the expression for the efficiency of a Carnot engine, sets an upper limit to the efficiency of a real engine such as a steam turbine. To maximize this upper limit and the actual efficiency of the real engine, the designer must make the intake temperature T_H as high as possible and the exhaust temperature T_C as low as possible (**Fig. 20.16**).

Figure **20.16** To maximize efficiency, the temperatures inside a jet engine are made as high as possible. Exotic ceramic materials are used that can withstand temperatures in excess of 1000°C without melting or becoming soft.

The exhaust temperature cannot be lower than the lowest temperature available for cooling the exhaust. For a steam turbine at an electric power plant, T_C may be the temperature of river or lake water; then we want the boiler temperature T_H to be as high as possible. The vapor pressures of all liquids increase rapidly with temperature, so we are limited by the mechanical strength of the boiler. At 500°C the vapor pressure of water is about 240×10^5 Pa (235 atm); this is about the maximum practical pressure in large present-day steam boilers.

The Kelvin Temperature Scale

In Chapter 17 we expressed the need for a temperature scale that doesn't depend on the properties of any particular material. We can now use the Carnot cycle to define such a scale. The thermal efficiency of a Carnot engine operating between two heat reservoirs at temperatures T_H and T_C is independent of the nature of the working substance and depends only on the temperatures. From Eq. (20.4), this thermal efficiency is

$$e = \frac{Q_H + Q_C}{Q_H} = 1 + \frac{Q_C}{Q_H}$$

Therefore the ratio Q_C/Q_H is the same for *all* Carnot engines operating between two given temperatures T_H and T_C.

Kelvin proposed that we *define* the ratio of the temperatures, T_C/T_H, to be equal to the magnitude of the ratio Q_C/Q_H of the quantities of heat absorbed and rejected:

$$\frac{T_C}{T_H} = \frac{|Q_C|}{|Q_H|} = -\frac{Q_C}{Q_H} \quad \text{(definition of Kelvin temperature)} \qquad (20.16)$$

Equation (20.16) looks identical to Eq. (20.13), but there is a subtle and crucial difference. The temperatures in Eq. (20.13) are based on an ideal-gas thermometer, as defined in Section 17.3, while Eq. (20.16) *defines* a temperature scale based on the Carnot cycle and the second law of thermodynamics and is independent of the behavior of any particular substance. Thus the **Kelvin temperature scale** is truly *absolute*. To complete the definition of the Kelvin scale, we assign, as in Section 17.3, the arbitrary value of 273.16 K to the temperature of the triple point of water. When a substance is taken around a Carnot cycle, the ratio of the heats absorbed and rejected, $|Q_H|/|Q_C|$, is equal to the ratio of the temperatures of the reservoirs *as expressed on the gas-thermometer scale* defined in Section 17.3. Since the triple point of water is chosen to be 273.16 K in both scales, it follows that *the Kelvin and ideal-gas scales are identical*.

The zero point on the Kelvin scale is called **absolute zero.** At absolute zero a system has its *minimum* possible total internal energy (kinetic plus potential). Because of quantum effects, however, it is *not* true that at $T = 0$, all molecular motion ceases. There are theoretical reasons for believing that absolute zero cannot be attained experimentally, although temperatures below 10^{-8} K have been achieved. The more closely we approach absolute zero, the more difficult it is to get closer. One statement of the *third law of thermodynamics* is that it is impossible to reach absolute zero in a finite number of thermodynamic steps.

CAUTION Changes to the Kelvin scale As we mentioned in Sections 17.3 and 18.3, in 2019 the Kelvin scale was redefined to be based not on the triple point of water but rather on the values of the joule ($1\ \text{J} = 1\ \text{kg}\ \text{kg} \cdot \text{m}^2/\text{s}^2$) and the Boltzmann constant. Despite this change in definition, the Kelvin and ideal-gas scales remain identical.

TEST YOUR UNDERSTANDING OF SECTION 20.6 An inventor looking for financial support comes to you with an idea for a gasoline engine that runs on a novel type of thermodynamic cycle. His design is made entirely of copper and is air-cooled. He claims that the engine will be 85% efficient. Should you invest in this marvelous new engine? (*Hint:* See Table 17.4.)

ANSWER

no The efficiency can be no better than that of a Carnot engine running between the same two temperature limits, $e_{\text{Carnot}} = 1 - (T_C/T_H)$ [Eq. (20.14)]. The temperature T_C of the cold reservoir for this air-cooled engine is about 300 K (ambient temperature), and the temperature T_H of the hot reservoir cannot exceed the melting point of copper, 1356 K (see Table 17.4). Hence the maximum possible Carnot efficiency is $e = 1 - (300\ \text{K})/(1356\ \text{K}) = 0.78$, or 78%. The temperature of any real engine would be less than this, so it would be impossible for the inventor's engine to attain 85% efficiency. You should invest your money elsewhere.

20.7 ENTROPY

The second law of thermodynamics, as we have stated it, is not an equation or a quantitative relationship but rather a statement of *impossibility*. However, the second law *can* be stated as a quantitative relationship with the concept of *entropy,* the subject of this section.

We have talked about several processes that proceed naturally in the direction of increasing randomness. Irreversible heat flow increases randomness: The molecules are initially sorted into hotter and cooler regions, but this sorting is lost when the system comes to thermal equilibrium. Adding heat to an object also increases average molecular speeds; therefore, molecular motion becomes more random. In the free expansion of a gas, the molecules have greater randomness of position after the expansion than before. **Figure 20.17** shows another process in which randomness increases.

Figure 20.17 When firecrackers explode, randomness increases: The neatly packaged chemicals within each firecracker are dispersed in all directions, and the stored chemical energy is converted to random kinetic energy of the fragments.

Entropy and Randomness

Entropy provides a *quantitative* measure of randomness. To introduce this concept, let's consider an infinitesimal isothermal expansion of an ideal gas. We add heat dQ and let the gas expand just enough to keep the temperature constant. Because the internal energy of an ideal gas depends on only its temperature, the internal energy is also constant; thus from the first law, the work dW done by the gas is equal to the heat dQ added. That is,

$$dQ = dW = p\,dV = \frac{nRT}{V}dV \qquad \text{so} \qquad \frac{dV}{V} = \frac{dQ}{nRT}$$

The gas is more disordered after the expansion than before: The molecules are moving in a larger volume and have more randomness of position. Thus the fractional volume change dV/V is a measure of the increase in randomness, and the above equation shows that it is proportional to the quantity dQ/T. We introduce the symbol S for the entropy of the system, and we define the infinitesimal entropy change dS during an infinitesimal reversible process at absolute temperature T as

$$dS = \frac{dQ}{T} \qquad \text{(infinitesimal reversible process)} \tag{20.17}$$

If a total amount of heat Q is added during a reversible isothermal process at absolute temperature T, the total entropy change $\Delta S = S_2 - S_1$ is given by

$$\Delta S = S_2 - S_1 = \frac{Q}{T} \qquad \text{(reversible isothermal process)} \tag{20.18}$$

Entropy has units of energy divided by temperature; the SI unit of entropy is 1 J/K.

We can see how the quotient Q/T is related to the increase in randomness. Higher temperature means greater randomness of motion. If the substance is initially cold, with little molecular motion, adding heat Q causes a substantial fractional increase in molecular motion and randomness. But if the substance is already hot, the same quantity of heat adds relatively little to the greater molecular motion already present. In the first case the quantity Q/T is large because T in the denominator is small; in the second case Q/T is small because T is large. So Q/T is an appropriate characterization of the increase in randomness when heat flows into a system.

CAUTION **Be mindful of the signs of Q and ΔS** When calculating the entropy change for a reversible process, remember that Q is positive if heat flows into a system and negative if heat flows out of a system. So the entropy of a system increases (ΔS is positive) when heat flows into it, and the entropy decreases (ΔS is negative) when heat flows out of it.

EXAMPLE 20.5 Entropy change in melting

What is the change of entropy of 1 kg of ice that is melted reversibly at 0°C and converted to water at 0°C? The heat of fusion of water is $L_f = 3.34 \times 10^5$ J/kg.

IDENTIFY and SET UP The melting occurs at a constant temperature $T = 0°\text{C} = 273$ K, so this is an *isothermal* reversible process. We can calculate the added heat Q required to melt the ice, then calculate the entropy change ΔS from Eq. (20.18).

EXECUTE The heat needed to melt the ice is $Q = mL_f = 3.34 \times 10^5$ J. Then from Eq. (20.18),

$$\Delta S = S_2 - S_1 = \frac{Q}{T} = \frac{3.34 \times 10^5\ \text{J}}{273\ \text{K}} = 1.22 \times 10^3\ \text{J/K}$$

EVALUATE This entropy increase corresponds to the increase in disorder when the water molecules go from the state of a crystalline solid to the much more randomly arranged state of a liquid. In *any* isothermal reversible process, the entropy change equals the heat transferred divided by the absolute temperature. When we refreeze the water, Q has the opposite sign, and the entropy change is $\Delta S = -1.22 \times 10^3$ J/K. The water molecules rearrange themselves into a crystal to form ice, so both randomness and entropy decrease.

KEYCONCEPT To find the entropy change of a system that goes through a reversible thermodynamic process at a constant temperature T, divide the heat Q that flows into the system during the process by T. If heat flows *in* (Q is positive), the entropy and disorder of the system *increase;* if heat flows *out* (Q is negative), the entropy and disorder of the system *decrease*.

Entropy in Reversible Processes

We can generalize the definition of entropy change to include *any* reversible process leading from one state to another, whether it is isothermal or not. We represent the process as a series of infinitesimal reversible steps. During a typical step, an infinitesimal quantity of heat dQ is added to the system at absolute temperature T. Then we sum (integrate) the quotients dQ/T for the entire process; that is,

$$\underbrace{\Delta S}_{\text{Entropy change in a reversible process}} = \int_{\underbrace{1}_{\text{Lower limit = initial state}}}^{\overbrace{2}^{\text{Upper limit = final state}}} \frac{\overbrace{dQ}^{\text{Infinitesimal heat flow into system}}}{\underbrace{T}_{\text{Absolute temperature}}} \tag{20.19}$$

Because entropy is a measure of the randomness of a system in any specific state, it must depend only on the current state of the system, not on its past history. (We'll verify this later.) When a system proceeds from an initial state with entropy S_1 to a final state with entropy S_2, the change in entropy $\Delta S = S_2 - S_1$ defined by Eq. (20.19) does not depend on the path leading from the initial to the final state but is the same for *all possible* processes leading from state 1 to state 2. Thus the entropy of a system must also have a definite value for any given state of the system. *Internal energy,* introduced in Chapter 19, also has this property, although entropy and internal energy are very different quantities.

Since entropy is a function only of the state of a system, we can also compute entropy changes in *irreversible* (nonequilibrium) processes for which Eqs. (20.17) and (20.19) are not applicable. We simply invent a path connecting the given initial and final states that *does* consist entirely of reversible equilibrium processes and compute the total entropy change for that path. It is not the actual path, but the entropy change must be the same as for the actual path.

As with internal energy, the above discussion does not tell us how to calculate entropy itself, but only the change in entropy in any given process. Just as with internal energy, we may arbitrarily assign a value to the entropy of a system in a specified reference state and then calculate the entropy of any other state with reference to this.

EXAMPLE 20.6 Entropy change in a temperature change

WITH VARIATION PROBLEMS

One kilogram of water at 0°C is heated to 100°C. Compute its change in entropy. Assume that the specific heat of water is constant at 4190 J/kg · K over this temperature range.

IDENTIFY and SET UP The entropy change of the water depends only on the initial and final states of the system, no matter whether the process is reversible or irreversible. We can imagine a reversible process in which the water temperature is increased in a sequence of infinitesimal steps dT. We can use Eq. (20.19) to integrate over all these steps and calculate the entropy change for such a reversible process. (Heating the water on a stove whose cooking surface is maintained at 100°C would be an irreversible process. The entropy change would be the same, however.)

EXECUTE From Eq. (17.14) the heat required to carry out each infinitesimal step is $dQ = mc\,dT$. Substituting this into Eq. (20.19) and integrating, we find

$$\Delta S = S_2 - S_1 = \int_1^2 \frac{dQ}{T} = \int_{T_1}^{T_2} mc\frac{dT}{T} = mc \ln\frac{T_2}{T_1}$$

$$= (1.00\text{ kg})(4190\text{ J/kg}\cdot\text{K})\left(\ln\frac{373\text{ K}}{273\text{ K}}\right) = 1.31 \times 10^3\text{ J/K}$$

EVALUATE The entropy change is positive, as it must be for a process in which the system absorbs heat. Our assumption about the specific heat is a pretty good one, since c for water varies by less than 1% between 0°C and 100°C (see Fig. 17.17).

CAUTION When $\Delta S = Q/T$ can (and cannot) be used In solving this problem you might be tempted to avoid doing an integral by using the simpler expression in Eq. (20.18), $\Delta S = Q/T$. This would be incorrect, however, because Eq. (20.18) is applicable only to *isothermal* processes, and the initial and final temperatures in our example are *not* the same. The *only* correct way to find the entropy change in a process with different initial and final temperatures is to use Eq. (20.19).

KEYCONCEPT The entropy change ΔS of a thermodynamic system does not depend on the path taken between the initial and final states of the system. You can calculate ΔS for any initial and final states by analyzing a *reversible* process that takes the system between these states; in general this involves evaluating an integral.

CONCEPTUAL EXAMPLE 20.7 Entropy change in a reversible adiabatic process

WITH VARIATION PROBLEMS

A gas expands adiabatically and reversibly. What is its change in entropy?

SOLUTION In an adiabatic process, no heat enters or leaves the system. Hence $dQ = 0$ and there is *no* change in entropy in this reversible process: $\Delta S = 0$. Every *reversible* adiabatic process is a constant-entropy process. (That's why such processes are also called *isentropic* processes.) The increase in randomness resulting from the gas occupying a greater volume is exactly balanced by the decrease in randomness associated with the lowered temperature and reduced molecular speeds.

KEYCONCEPT In the special case where a system undergoes a reversible adiabatic process, the system undergoes *no* change in entropy.

EXAMPLE 20.8 Entropy change in a free expansion

WITH VARIATION PROBLEMS

A partition divides a thermally insulated box into two compartments, each of volume V (**Fig. 20.18**). Initially, one compartment contains n moles of an ideal gas at temperature T, and the other compartment is evacuated. We break the partition and the gas expands, filling both compartments. What is the entropy change in this free-expansion process?

IDENTIFY and SET UP For this process, $Q = 0$, $W = 0$, $\Delta U = 0$, and therefore (because the system is an ideal gas) $\Delta T = 0$. We might think that the entropy change is zero because there is no heat exchange. But Eq. (20.19) can be used to calculate entropy changes for *reversible* processes only; this free expansion is *not* reversible, and there *is* an entropy change. As we mentioned at the beginning of this section, entropy increases in a free expansion because the positions of the molecules are more random than before the expansion. To calculate ΔS, we recall that the entropy change depends only on the initial and final states. We can devise a *reversible* process having the same endpoints as this free expansion, and in general we can then use Eq. (20.19) to calculate its entropy change, which will be the same as for the free expansion. An appropriate reversible process is an *isothermal* expansion from V to $2V$ at temperature T, which allows us to use the simpler Eq. (20.18) to calculate ΔS. The gas does work W during this expansion, so an equal amount of heat Q must be supplied to keep the internal energy constant.

EXECUTE We saw in Example 19.1 that the work done by n moles of ideal gas in an isothermal expansion from V_1 to V_2 is $W = nRT\ln(V_2/V_1)$. With $V_1 = V$ and $V_2 = 2V$, we have

$$Q = W = nRT\ln\frac{2V}{V} = nRT\ln 2$$

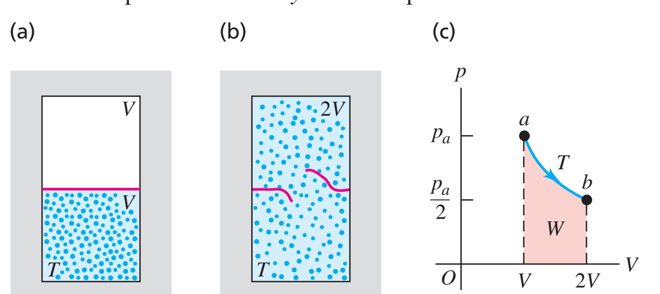
From Eq. (20.18), the entropy change is

$$\Delta S = \frac{Q}{T} = nR\ln 2$$

EVALUATE For 1 mole, $\Delta S = (1\ \text{mol})(8.314\ \text{J/mol}\cdot\text{K})(\ln 2) = 5.76\ \text{J/K}$. The entropy change is positive, as we predicted. The factor $(\ln 2)$ in our answer is a result of the volume having increased by a factor of 2, from V to $2V$. Can you show that if the volume increases in a free expansion from V to xV, where x is an arbitrary number, the entropy change is $\Delta S = nR\ln x$?

KEYCONCEPT To calculate the entropy change of a system in an *irreversible* process such as a free expansion, you must consider a *reversible* process that takes the system from the same initial state to the same final state.

Figure **20.18** (a, b) Free expansion of an insulated ideal gas. (c) The free-expansion process doesn't pass through equilibrium states from a to b. However, the entropy change $S_b - S_a$ can be calculated by using the isothermal path shown or *any* reversible path from a to b.



EXAMPLE 20.9 Entropy and the Carnot cycle

WITH VARIATION PROBLEMS

For the Carnot engine in Example 20.2 (Section 20.6), what is the total entropy change during one cycle?

IDENTIFY and SET UP All four steps in the Carnot cycle (see Fig. 20.13) are reversible, so we can use our expressions for the entropy change ΔS in a reversible process. We find ΔS for each step and add them to get ΔS for the complete cycle.

EXECUTE There is no entropy change during the adiabatic expansion $b \rightarrow c$ or the adiabatic compression $d \rightarrow a$. During the isothermal expansion $a \rightarrow b$ at $T_H = 500$ K, the engine takes in 2000 J of heat, and from Eq. (20.18),

$$\Delta S_H = \frac{Q_H}{T_H} = \frac{2000\ \text{J}}{500\ \text{K}} = 4.0\ \text{J/K}$$

During the isothermal compression $c \rightarrow d$ at $T_C = 350$ K, the engine gives off 1400 J of heat, and

$$\Delta S_C = \frac{Q_C}{T_C} = \frac{-1400\ \text{J}}{350\ \text{K}} = -4.0\ \text{J/K}$$

The total entropy change in the engine during one cycle is $\Delta S_{tot} = \Delta S_H + \Delta S_C = 4.0\ \text{J/K} + (-4.0\ \text{J/K}) = 0$.

EVALUATE The result $\Delta S_{total} = 0$ tells us that when the Carnot engine completes a cycle, it has the same entropy as it did at the beginning of the cycle. We'll explore this result in the next subsection.

What is the total entropy change of the engine's *environment* during this cycle? During the reversible isothermal expansion $a \rightarrow b$, the hot (500 K) reservoir gives off 2000 J of heat, so its entropy change is $(-2000\ \mathrm{J})/(500\ \mathrm{K}) = -4.0\ \mathrm{J/K}$. During the reversible isothermal compression $c \rightarrow d$, the cold (350 K) reservoir absorbs 1400 J of heat, so its entropy change is $(+1400\ \mathrm{J})/(350\ \mathrm{K}) = +4.0\ \mathrm{J/K}$. Thus the hot and cold reservoirs each have an entropy change, but the sum of these changes—that is, the total entropy change of the system's environment—is zero.

These results apply to the special case of the Carnot cycle, for which *all* of the processes are reversible. In this case the total entropy change of the system and the environment together is zero. We'll see that if the cycle includes irreversible processes (as is the case for the Otto and Diesel cycles of Section 20.3), the total entropy change of the system and the environment *cannot* be zero, but rather must be positive.

KEYCONCEPT For the special case of a heat engine that uses the Carnot cycle, the net entropy change of the engine and its environment (the hot and cold reservoirs) in a complete cycle is zero. This is not the case for other, less ideal cycles.

Entropy in Cyclic Processes

Example 20.9 showed that the total entropy change for a cycle of a particular Carnot engine, which uses an ideal gas as its working substance, is zero. This result follows directly from Eq. (20.13), which we can rewrite as

$$\frac{Q_H}{T_H} + \frac{Q_C}{T_C} = 0 \qquad (20.20)$$

The quotient Q_H/T_H equals ΔS_H, the entropy change of the engine that occurs at $T = T_H$. Likewise, Q_C/T_C equals ΔS_C, the (negative) entropy change of the engine that occurs at $T = T_C$. Hence Eq. (20.20) says that $\Delta S_H + \Delta S_C = 0$; that is, there is zero net entropy change in one cycle.

What about Carnot engines that use a different working substance? According to the second law, *any* Carnot engine operating between given temperatures T_H and T_C has the same efficiency $e = 1 - T_C/T_H$ [Eq. (20.14)]. Combining this expression for e with Eq. (20.4), $e = 1 + Q_C/Q_H$, just reproduces Eq. (20.20). So Eq. (20.20) is valid for any Carnot engine working between these temperatures, whether its working substance is an ideal gas or not. We conclude that *the total entropy change in one cycle of any Carnot engine is zero.*

This result can be generalized to show that the total entropy change during *any* reversible cyclic process is zero. A reversible cyclic process appears on a pV-diagram as a closed path (**Fig. 20.19a**). We can approximate such a path as closely as we like by a sequence of isothermal and adiabatic processes forming parts of many long, thin Carnot cycles (Fig. 20.19b). The total entropy change for the full cycle is the sum of the entropy changes for each small Carnot cycle, each of which is zero. So **the total entropy change during *any* reversible cycle is zero:**

$$\int \frac{dQ}{T} = 0 \qquad \text{(reversible cyclic process)} \qquad (20.21)$$

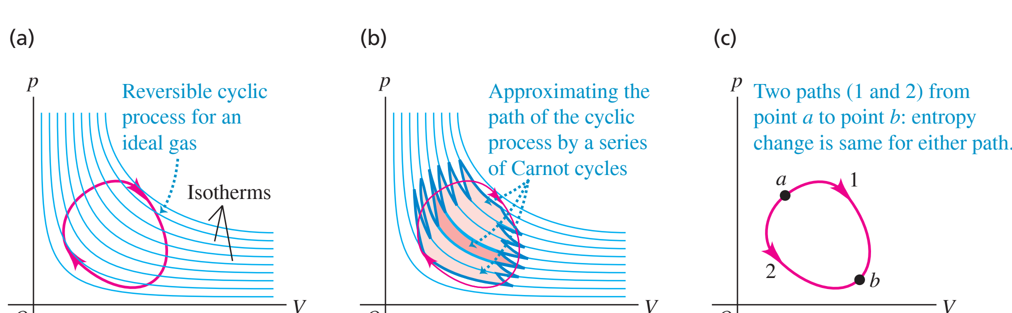


Figure **20.19** (a) A reversible cyclic process for an ideal gas is shown as a red closed path on a pV-diagram. Several ideal-gas isotherms are shown in blue. (b) We can approximate the path in (a) by a series of long, thin Carnot cycles; one of these is highlighted in color. The total entropy change is zero for each Carnot cycle and for the actual cyclic process. (c) The entropy change between points a and b is independent of the path.

BIO APPLICATION Entropy Changes in a Living Organism When a puppy or other growing animal eats, it takes organized chemical energy from the food and uses it to make new cells that are even more highly organized. This process alone lowers entropy. But most of the energy in the food is either excreted in the animal's feces or used to generate heat, processes that lead to a large increase in entropy. So while the entropy of the animal alone decreases, the *total* entropy of animal plus food *increases*.

It follows that when a system undergoes a reversible process leading from any state a to any other state b, *the entropy change of the system is independent of the path* (Fig. 20.19c). If the entropy change for path 1 were different from the change for path 2, the system could be taken along path 1 and then backward along path 2 to the starting point, with a nonzero net change in entropy. This would violate the conclusion that the total entropy change in such a cyclic process must be zero. Because the entropy change in such processes is independent of path, we conclude that in any given state, the system has a definite value of entropy that depends only on the state, not on the processes that led to that state.

Entropy in Irreversible Processes

In an idealized, reversible process involving only equilibrium states, the total entropy change of the system and its surroundings is zero. But all *irreversible* processes involve an increase in entropy. Unlike energy, *entropy is not a conserved quantity.* The entropy of an isolated system *can* change, but as we shall see, it can never decrease. The free expansion of a gas (Example 20.8) is an irreversible process in an isolated system in which there is an entropy increase.

EXAMPLE 20.10 Entropy change in an irreversible process — WITH VARIATION PROBLEMS

Suppose 1.00 kg of water at 100°C is placed in thermal contact with 1.00 kg of water at 0°C. What is the total change in entropy? Assume that the specific heat of water is constant at 4190 J/kg · K over this temperature range.

IDENTIFY and SET UP This process involves irreversible heat flow because of the temperature differences. There are equal masses of 0°C water and 100°C water, so the final temperature is the average of these two temperatures: 50°C = 323 K. Although the processes are irreversible, we can calculate the entropy changes for the (initially) hot water and the (initially) cold water by assuming that the process occurs reversibly. As in Example 20.6, we must use Eq. (20.19) to calculate ΔS for each substance because the temperatures are not constant.

EXECUTE The entropy changes of the hot water (subscript H) and the cold water (subscript C) are

$$\Delta S_{\mathrm{H}} = mc\int_{T_1}^{T_2}\frac{dT}{T} = (1.00\text{ kg})(4190\text{ J/kg}\cdot\text{K})\int_{373\text{ K}}^{323\text{ K}}\frac{dT}{T}$$

$$= (4190\text{ J/K})\left(\ln\frac{323\text{ K}}{373\text{ K}}\right) = -603\text{ J/K}$$

$$\Delta S_{\mathrm{C}} = (4190\text{ J/K})\left(\ln\frac{323\text{ K}}{273\text{ K}}\right) = +705\text{ J/K}$$

The *total* entropy change of the system is

$$\Delta S_{\mathrm{tot}} = \Delta S_{\mathrm{H}} + \Delta S_{\mathrm{C}} = (-603\text{ J/K}) + 705\text{ J/K} = +102\text{ J/K}$$

EVALUATE An irreversible heat flow in an isolated system is accompanied by an increase in entropy. We could reach the same end state by mixing the hot and cold water, which is also an irreversible process; the total entropy change, which depends only on the initial and final states of the system, would again be 102 J/K.

Note that the entropy of the system increases *continuously* as the two quantities of water come to equilibrium. For example, the first 4190 J of heat transferred cools the hot water to 99°C and warms the cold water to 1°C. The net change in entropy for this step is approximately

$$\Delta S = \frac{-4190\text{ J}}{373\text{ K}} + \frac{4190\text{ J}}{273\text{ K}} = +4.1\text{ J/K}$$

Can you show in a similar way that the net entropy change is positive for *any* one-degree temperature change leading to the equilibrium condition?

KEYCONCEPT When heat flows irreversibly from an object at higher temperature to one at lower temperature, the entropy of the high-temperature object decreases and the entropy of the low-temperature object increases. The *net* entropy of the two objects together always increases.

Entropy and the Second Law

The results of Example 20.10 about the flow of heat from a higher to a lower temperature are characteristic of *all* natural (that is, irreversible) processes. When we include the entropy changes of all the systems taking part in the process, the increases in entropy are always greater than the decreases. In the special case of a *reversible* process, the increases

Figure **20.20** The mixing of colored ink and water starts from a state of low entropy in which each fluid is separate and distinct from the other. In the final state, both the ink and water molecules are spread randomly throughout the volume of liquid, so the entropy is greater. Spontaneous unmixing of the ink and water, a process in which there would be a net decrease in entropy, is never observed.

and decreases are equal. Hence we can state the general principle: **When all systems taking part in a process are included, the entropy either remains constant or increases.** This gives us a third, alternative way of stating the second law of thermodynamics:

SECOND LAW OF THERMODYNAMICS ("entropy" statement) No process is possible in which the total entropy decreases, when all systems that take part in the process are included.

This is equivalent to the "engine" and "refrigerator" statements that we discussed in Section 20.5. **Figure 20.20** shows a specific example of this general principle.

The increase of entropy in every natural, irreversible process measures the increase of randomness in the universe associated with that process. Consider again the example of mixing hot and cold water (Example 20.10). We *might* have used the hot and cold water as the high- and low-temperature reservoirs of a heat engine. While removing heat from the hot water and giving heat to the cold water, we could have obtained some mechanical work. But once the hot and cold water have been mixed and have come to a uniform temperature, this opportunity to convert heat to mechanical work is lost irretrievably. The lukewarm water will never *unmix* itself and separate into hotter and colder portions. No decrease in *energy* occurs when the hot and cold water are mixed. What has been lost is the *opportunity* to convert part of the heat from the hot water into mechanical work. Hence when entropy increases, energy becomes less *available,* and the universe becomes more random or "run down."

TEST YOUR UNDERSTANDING OF SECTION 20.7 Suppose 2.00 kg of water at 50°C spontaneously changes temperature, so that half of the water cools to 0°C while the other half spontaneously warms to 100°C. (All of the water remains liquid, so it doesn't freeze or boil.) What would be the entropy change of the water? Is this process possible? (*Hint:* See Example 20.10.)

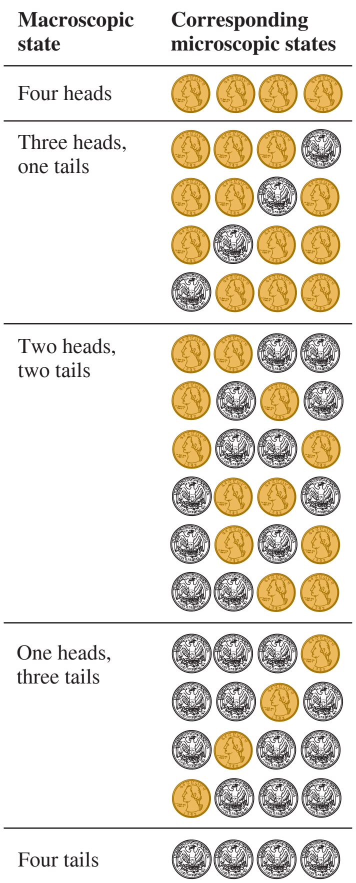
ANSWER

−102 J/K, no The process described is exactly the opposite of the process used in Example 20.10. The result violates the second law of thermodynamics, which states that the entropy of an isolated system cannot decrease.

20.8 MICROSCOPIC INTERPRETATION OF ENTROPY

We described in Section 19.4 how the internal energy of a system could be calculated, at least in principle, by adding up all the kinetic energies of its constituent particles and all the potential energies of interaction among the particles. This is called a *microscopic* calculation of the internal energy. We can also make a microscopic calculation of the entropy S of a system. Unlike energy, however, entropy is not something that belongs to each individual particle or pair of particles in the system. Rather, entropy is a measure of the randomness of the system as a whole. To see how to calculate entropy microscopically, we first have to introduce the idea of *macroscopic* and *microscopic states.*

Figure **20.21** All possible microscopic states of four coins. There can be several possible microscopic states for each macroscopic state.



Suppose you toss N identical coins on the floor, and half of them show heads and half show tails. This is a description of the large-scale or **macroscopic state** of the system of N coins. A description of the **microscopic state** of the system includes information about each individual coin: Coin 1 was heads, coin 2 was tails, coin 3 was tails, and so on. There can be many microscopic states that correspond to the same macroscopic description. For instance, with $N = 4$ coins there are six possible states in which half are heads and half are tails (**Fig. 20.21**). The number of microscopic states grows rapidly with increasing N; for $N = 100$ there are $2^{100} = 1.27 \times 10^{30}$ microscopic states, of which 1.01×10^{29} are half heads and half tails.

The least probable outcomes of the coin toss are the states that are either all heads or all tails. It is certainly possible that you could throw 100 heads in a row, but don't bet on it; the probability of doing this is only 1 in 1.27×10^{30}. The most probable outcome of tossing N coins is that half are heads and half are tails. The reason is that this *macroscopic* state has the greatest number of corresponding *microscopic* states, as Fig. 20.21 shows.

To make the connection to the concept of entropy, note that the macroscopic description "all heads" completely specifies the state of each one of the N coins. The same is true if the coins are all tails. But the macroscopic description "half heads, half tails" by itself tells you very little about the state (heads or tails) of each individual coin. Compared to the state "all heads" or "all tails," the state "half heads, half tails" has much greater *randomness* because the system could be in any of a much greater number of possible microscopic states. Hence the "half heads, half tails" state has much greater entropy (which is a quantitative measure of randomness).

Now instead of N coins, consider a mole of an ideal gas containing Avogadro's number of molecules. The macroscopic state of this gas is given by its pressure p, volume V, and temperature T; a description of the microscopic state involves stating the position and velocity for each molecule in the gas. At a given pressure, volume, and temperature, the gas may be in any one of an astronomically large number of microscopic states, depending on the positions and velocities of its 6.02×10^{23} molecules. If the gas undergoes a free expansion into a greater volume, the range of possible positions increases, as does the number of possible microscopic states. The system becomes more random, and the entropy increases as calculated in Example 20.8 (Section 20.7).

We can draw the following general conclusion: **For any thermodynamic system, the most probable macroscopic state is the one with the greatest number of corresponding microscopic states, which is also the macroscopic state with the greatest randomness and the greatest entropy.**

Calculating Entropy: Microscopic States

Let w represent the number of possible microscopic states for a given macroscopic state. (For the four coins shown in Fig. 20.21 the state of four heads has $w = 1$, the state of three heads and one tails has $w = 4$, and so on.) Then the entropy S of a macroscopic state can be shown to be given by

Expression for entropy in microscopic terms

$$S = k \ln w \qquad (20.22)$$

Number of microscopic states for the given macroscopic state

Boltzmann constant (gas constant per molecule)

(We introduced the Boltzmann constant in Section 18.3.) As Eq. (20.22) shows, increasing the number of possible microscopic states w increases the entropy S.

What matters in a thermodynamic process is not the absolute entropy S but the *difference* in entropy between the initial and final states. Hence an equally valid and useful definition would be $S = k \ln w + C$, where C is a constant, since C cancels in any calculation of an entropy difference between two states. But it's convenient to set this constant equal to zero and use Eq. (20.22). With this choice, since the smallest possible value of w is unity, the smallest possible value of S for any system is $k \ln 1 = 0$. Entropy can *never* be negative.

In practice, calculating w is a difficult task, so Eq. (20.22) is typically used only to calculate the absolute entropy S of certain special systems. But we can use this relationship to calculate *differences* in entropy between one state and another. Consider a system that undergoes a thermodynamic process that takes it from macroscopic state 1, for which there are w_1 possible microscopic states, to macroscopic state 2, with w_2 associated microscopic states. The change in entropy in this process is

$$\Delta S = S_2 - S_1 = k \ln w_2 - k \ln w_1 = k \ln \frac{w_2}{w_1} \qquad (20.23)$$

The *difference* in entropy between the two macroscopic states depends on the *ratio* of the numbers of possible microscopic states.

As the following example shows, using Eq. (20.23) to calculate a change in entropy from one macroscopic state to another gives the same results as considering a reversible process connecting those two states and using Eq. (20.19).

EXAMPLE 20.11 A microscopic calculation of entropy change

Use Eq. (20.23) to calculate the entropy change in the free expansion of n moles of gas at temperature T described in Example 20.8 (**Fig. 20.22**).

IDENTIFY and SET UP We are asked to calculate the entropy change by using the number of microscopic states in the initial and final macroscopic states (Figs. 20.22a and b). When the partition is broken, no work is done, so the velocities of the molecules are unaffected. But each molecule now has twice as much volume in which it can move and hence has twice the number of possible positions. This is all we need to calculate the entropy change using Eq. (20.23).

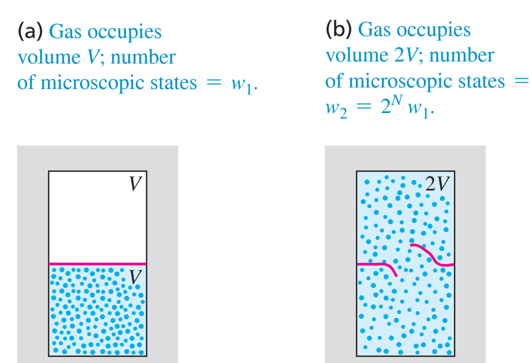
EXECUTE Let w_1 be the number of microscopic states of the system as a whole when the gas occupies volume V (Fig. 20.22a). The number of molecules is $N = nN_A$, and each of these N molecules has twice as many possible states after the partition is broken. Hence the number w_2 of microscopic states when the gas occupies volume $2V$ (Fig. 20.22b) is greater by a factor of 2^N; that is, $w_2 = 2^N w_1$. The change in entropy in this process is

$$\Delta S = k \ln \frac{w_2}{w_1} = k \ln \frac{2^N w_1}{w_1}$$
$$= k \ln 2^N = Nk \ln 2$$

Since $N = nN_A$ and $k = R/N_A$, this becomes

$$\Delta S = (nN_A)(R/N_A) \ln 2$$
$$= nR \ln 2$$

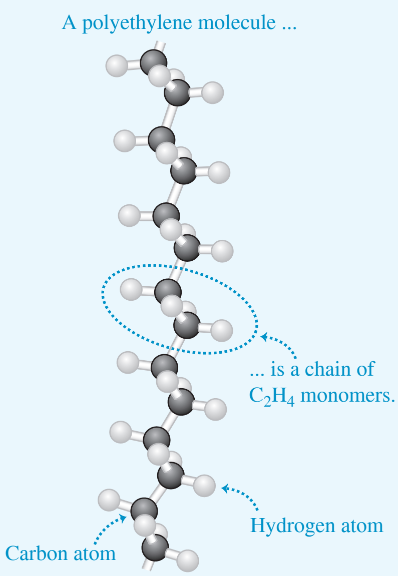
Figure **20.22** In a free expansion of N molecules in which the volume doubles, the number of possible microscopic states increases by a factor of 2^N.

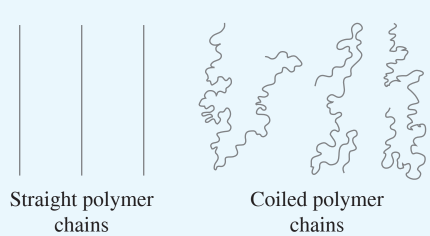


EVALUATE We found the same result as in Example 20.8, but without any reference to the thermodynamic path taken.

KEYCONCEPT An alternative way to calculate entropy changes is in terms of the number of microscopic states available to the system. If there are w possible microscopic states of the system for a given macroscopic state, the entropy for the macroscopic state equals the Boltzmann constant k multiplied by the natural logarithm of w.

APPLICATION Polymers Coil in Solution A molecule of polyethylene, the most common plastic, is a polymer—a long chain of monomer units (C_2H_4). In solution these molecules coil on themselves, and the entropy concept explains why. The polymer can coil in many ways (microscopic states), but there is only one microscopic state in which the polymer is fully stretched out. Thus the entropy of the coiled polymer is much greater than that of a stretched-out polymer. The second law of thermodynamics says that isolated systems always move toward greater entropy, so we expect a polymer chain in solution to be in a coiled state.





Microscopic States and the Second Law

The relationship between entropy and the number of microscopic states gives us new insight into the entropy statement of the second law of thermodynamics: that the entropy of a closed system can never decrease. From Eq. (20.22) this means that a closed system can never spontaneously undergo a process that decreases the number of possible microscopic states.

An example of such a forbidden process would be if all of the air in your room spontaneously moved to one half of the room, leaving a vacuum in the other half. Such a "free compression" would be the reverse of the free expansion of Examples 20.8 and 20.11. This would decrease the number of possible microscopic states by a factor of 2^N. Strictly speaking, this process is not impossible! The probability of finding a given molecule in one half of the room is $\frac{1}{2}$, so the probability of finding all of the molecules in one half of the room at once is $\left(\frac{1}{2}\right)^N$. (This is exactly the same as the probability of having a tossed coin come up heads N times in a row.) This probability is *not* zero. But lest you worry about suddenly finding yourself gasping for breath in the evacuated half of your room, consider that a typical room might hold 1000 moles of air, and so $N = 1000N_A = 6.02 \times 10^{26}$ molecules. The probability of all the molecules being in the same half of the room is therefore $\left(\frac{1}{2}\right)^{6.02\times10^{26}}$. Expressed as a decimal, this number has more than 10^{26} zeros to the right of the decimal point!

Because the probability of such a "free compression" taking place is so vanishingly small, it has almost certainly never occurred anywhere in the universe since the beginning of time. We conclude that for all practical purposes the second law of thermodynamics is never violated.

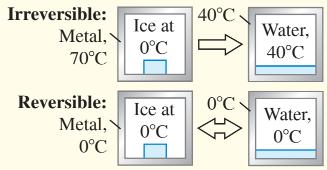
TEST YOUR UNDERSTANDING OF SECTION 20.8 A quantity of N molecules of an ideal gas initially occupies volume V. The gas then expands to volume $2V$. The number of microscopic states of the gas increases in this expansion. Under which of the following circumstances will this number increase the most? (i) If the expansion is reversible and isothermal; (ii) if the expansion is reversible and adiabatic; (iii) the number will change by the same amount for both circumstances.

ANSWER

(i) For case (i), we saw in Example 20.8 (Section 20.7) that for an ideal gas, the entropy change in a free expansion is the same as in an isothermal expansion. From Eq. (20.23), this implies that the ratio of the number of microscopic states after and before the expansion, w_2/w_1, is also the same for these two cases. From Example 20.11, $w_2/w_1 = 2^N$, so the number of microscopic states increases by a factor 2^N. For case (ii), in a reversible expansion the entropy change is $\Delta S = \int dQ/T = 0$; if the expansion is adiabatic there is no heat flow, so $\Delta S = 0$. From Eq. (20.23), $w_2/w_1 = 1$ and there is *no* change in the number of microscopic states. The difference is that in an adiabatic expansion the temperature drops and the molecules move more slowly, so they have fewer microscopic states available to them than in an isothermal expansion.

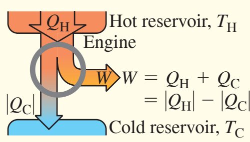
CHAPTER 20 SUMMARY

Reversible and irreversible processes: A reversible process is one whose direction can be reversed by an infinitesimal change in the conditions of the process, and in which the system is always in or very close to thermal equilibrium. All other thermodynamic processes are irreversible.



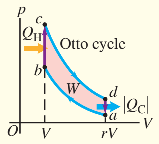
Heat engines: A heat engine takes heat Q_H from a source, converts part of it to work W, and discards the remainder $|Q_C|$ at a lower temperature. The thermal efficiency e of a heat engine measures how much of the absorbed heat is converted to work. (See Example 20.1.)

$$e = \frac{W}{Q_H} = 1 + \frac{Q_C}{Q_H} = 1 - \left|\frac{Q_C}{Q_H}\right| \qquad (20.4)$$



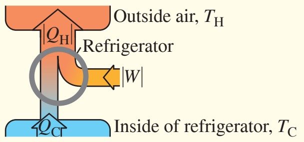
The Otto cycle: A gasoline engine operating on the Otto cycle has a theoretical maximum thermal efficiency e that depends on the compression ratio r and the ratio of heat capacities γ of the working substance.

$$e = 1 - \frac{1}{r^{\gamma-1}} \qquad (20.6)$$

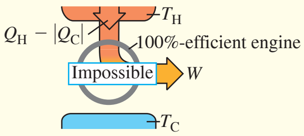


Refrigerators: A refrigerator takes heat Q_C from a colder place, has a work input $|W|$, and discards heat $|Q_H|$ at a warmer place. The effectiveness of the refrigerator is given by its coefficient of performance K.

$$K = \frac{|Q_C|}{|W|} = \frac{|Q_C|}{|Q_H| - |Q_C|} \qquad (20.9)$$



The second law of thermodynamics: The second law of thermodynamics describes the directionality of natural thermodynamic processes. It can be stated in two equivalent forms. The *engine* statement is that no cyclic process can convert heat completely into work. The *refrigerator* statement is that no cyclic process can transfer heat from a colder place to a hotter place with no input of mechanical work.

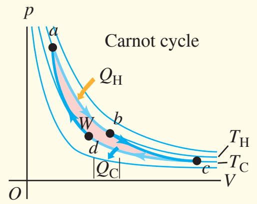


The Carnot cycle: The Carnot cycle operates between two heat reservoirs at temperatures T_H and T_C and uses only reversible processes. Its thermal efficiency depends only on T_H and T_C. An additional equivalent statement of the second law is that no engine operating between the same two temperatures can be more efficient than a Carnot engine. (See Examples 20.2
and 20.3.)

$$e_{Carnot} = 1 - \frac{T_C}{T_H} = \frac{T_H - T_C}{T_H} \qquad (20.14)$$

A Carnot engine run backward is a Carnot refrigerator. Its coefficient of performance depends on only T_H and T_C. Another form of the second law states that no refrigerator operating between the same two temperatures can have a larger coefficient of performance than a Carnot refrigerator. (See Example 20.4.)

$$K_{Carnot} = \frac{T_C}{T_H - T_C} \qquad (20.15)$$



Entropy: Entropy is a quantitative measure of the randomness of a system. The entropy change in any reversible process depends on the amount of heat flow and the absolute temperature T. Entropy depends only on the state of the system, and the change in entropy between given initial and final states is the same for all processes leading from one state to the other. This fact can be used to find the entropy change in an irreversible process. (See Examples 20.5–20.10.)

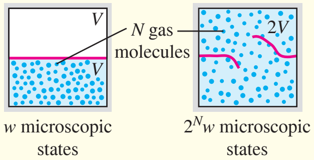
An important statement of the second law of thermodynamics is that the entropy of an isolated system may increase but can never decrease. When a system interacts with its surroundings, the total entropy change of system and surroundings can never decrease. When the interaction involves only reversible processes, the total entropy is constant and $\Delta S = 0$; when there is any irreversible process, the total entropy increases and $\Delta S > 0$.

$$\Delta S = \int_1^2 \frac{dQ}{T} \quad (20.19)$$

(reversible process)

Entropy and microscopic states: When a system is in a particular macroscopic state, the particles that make up the system may be in any of w possible microscopic states. The greater the number w, the greater the entropy. (See Example 20.11.)

$$S = k \ln w \quad (20.22)$$



Chapter 20 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLE 20.1 (Section 20.2) before attempting these problems.

VP20.1.1 A diesel engine has efficiency 0.180. (a) In order for this engine to do 1.24×10^4 J of work, how many joules of heat must it take in? (b) How many joules of this heat is discarded?

VP20.1.2 In one cycle a heat engine absorbs 3.82×10^4 J of heat from the hot reservoir and rejects 3.16×10^4 J of heat to the cold reservoir. What is the efficiency of this engine?

VP20.1.3 Measurements of a gasoline engine show that it has an efficiency of 0.196 and that it exhausts 4.96×10^8 J of heat during 20 minutes of operation. During that time, (a) how much heat does the engine take in and (b) how much work does the engine do?

VP20.1.4 An aircraft piston engine that burns gasoline (heat of combustion 5.0×10^7 J/kg) has a power output of 1.10×10^5 W. (a) How much work does this engine do in 1.00 h? (b) This engine burns 34 kg of gasoline per hour. How much heat does the engine take in per hour? (c) What is the efficiency of the engine?

Be sure to review EXAMPLES 20.2, 20.3 and 20.4 (Section 20.6) before attempting these problems.

VP20.4.1 In one cycle a Carnot engine takes in 8.00×10^4 J of heat and does 1.68×10^4 J of work. The temperature of the engine's cold reservoir is 25.0°C. (a) What is the efficiency of this engine? (b) How much heat does this engine exhaust per cycle? (c) What is the temperature (in °C) of the hot reservoir?

VP20.4.2 For the Carnot cycle described in Example 20.3, you change the temperature of the cold reservoir from 27°C to −73°C. The initial pressure and volume at point a are unchanged, the volume still doubles during the isothermal expansion $a \rightarrow b$, and the volume still decreases by one-half during the isothermal compression $c \rightarrow d$. For this modified cycle, calculate (a) the new efficiency of the cycle and (b) the amount of work done in each of the four steps of the cycle.

VP20.4.3 A Carnot refrigerator has a cold reservoir at −10.0°C and a hot reservoir at 25.0°C. (a) What is its coefficient of performance? (b) How much work input does this refrigerator require to remove 4.00×10^6 J of heat from the cold reservoir?

VP20.4.4 A Carnot engine uses the expansion and compression of n moles of argon gas, for which $C_V = \frac{3}{2}R$. This engine operates between temperatures T_C and T_H. During the isothermal expansion $a \rightarrow b$, the volume of the gas increases from V_a to $V_b = 2V_a$. (a) Calculate the work W_{ab} done during the isothermal expansion $a \rightarrow b$. Give your answer in terms of n, R, and T_H. (b) Calculate the work W_{bc} done during the adiabatic expansion $b \rightarrow c$. Give your answer in terms of n, R, T_C and T_H. (c) For this engine, $W_{ab} = W_{bc}$. Find the ratio T_C/T_H and the efficiency of the engine.

Be sure to review EXAMPLES 20.6, 20.7, 20.8, 20.9, and 20.10 (Section 20.7) before attempting these problems.

VP20.10.1 Ethanol melts at 159 K (heat of fusion 1.042×10^5 J/kg) and boils at 351 K (heat of vaporization 8.54×10^5 J/kg). Liquid ethanol has a specific heat of 2428 J/kg · K (which we assume does not depend on temperature). If you have 1.00 kg of ethanol originally in the solid state at 159 K, calculate the change in entropy of the ethanol when it (a) melts at 159 K, (b) increases in temperature as a liquid from 159 K to 351 K, and (c) boils at 351 K.

VP20.10.2 Initially 5.00 mol of helium (which we can treat as an ideal gas) occupies volume 0.120 m^3 and is at temperature 20.0°C. You allow the helium to expand so that its final volume is 0.360 m^3 and its final temperature is also 20.0°C. Calculate the net change in entropy of the helium if (a) you make the helium expand isothermally from 0.120 m^3 to 0.360 m^3 and (b) you first increase the temperature of the helium so that it expands at constant pressure from 0.120 m^3 to 0.360 m^3, then cool the helium at constant volume to 20.0°C.

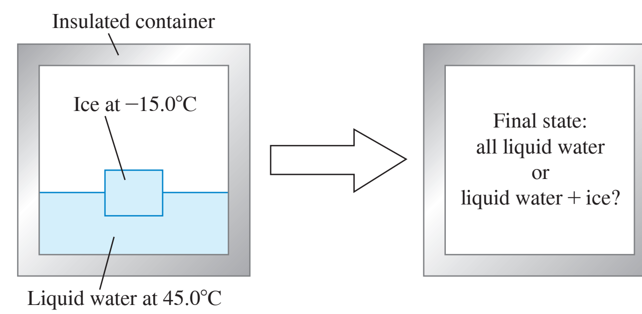
VP20.10.3 In one cycle of operation, an engine that does *not* use the Carnot cycle takes in 8.00×10^4 J of heat from a reservoir at 260°C, does 1.60×10^4 J of work, and rejects 6.40×10^4 J of heat to a reservoir at 20°C. At the end of the cycle, the engine is in the same state as at the beginning of the cycle. Calculate the change in entropy in one cycle for (a) the engine, (b) the reservoir at 260°C, (c) the reservoir at 20°C, and (d) the system of engine and two reservoirs. Does the net entropy of the system increase, decrease, or stay the same?

VP20.10.4 You combine 1.00 kg of ice (heat of fusion 3.34×10^5 J/K) at 0.0°C and 0.839 kg of liquid water (specific heat 4.19×10^3 J/K)) at 95.0°C. When the system comes to equilibrium, all of the ice has melted and the temperature of the mixture is 0.0°C. Calculate the change in entropy for (a) the ice as it melts, (b) the water that was initially at 95.0°C, and (c) the combination of the two. Does the net entropy of the system increase, decrease, or stay the same?

BRIDGING PROBLEM Entropy Changes: Cold Ice in Hot Water

An insulated container of negligible mass holds 0.600 kg of water at 45.0°C. You put a 0.0500 kg ice cube at −15.0°C in the water (**Fig. 20.23**). (a) Calculate the final temperature of the water once the ice has melted. (b) Calculate the change in entropy of the system.

Figure **20.23** What becomes of this ice–water mixture?



SOLUTION GUIDE

IDENTIFY and SET UP

1. Make a list of the known and unknown quantities, and identify the target variables.
2. How will you find the final temperature of the ice–water mixture? How will you decide whether or not all the ice melts?
3. Once you find the final temperature of the mixture, how will you determine the changes in entropy of (i) the ice initially at −15.0°C and (ii) the water initially at 45.0°C?

EXECUTE

4. Use the methods of Chapter 17 to calculate the final temperature T. (*Hint:* First assume that all of the ice melts, then write an equation which says that the heat that flows into the ice equals the heat that flows out of the water. If your assumption is correct, the final temperature that you calculate will be greater than 0°C. If your assumption is incorrect, the final temperature will be 0°C or less, which means that some ice remains. You'll then need to redo the calculation to account for this.)
5. Use your result from step 4 to calculate the entropy changes of the ice and the water. (*Hint:* You must include the heat flow associated with temperature changes, as in Example 20.6, as well as the heat flow associated with the change of phase.)
6. Find the total change in entropy of the system.

EVALUATE

7. Do the signs of the entropy changes make sense? Why or why not?

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q20.1 A pot is half-filled with water, and a lid is placed on it, forming a tight seal so that no water vapor can escape. The pot is heated on a stove, forming water vapor inside the pot. The heat is then turned off and the water vapor condenses back to liquid. Is this cycle reversible or irreversible? Why?

Q20.2 Give two examples of reversible processes and two examples of irreversible processes in purely mechanical systems, such as blocks sliding on planes, springs, pulleys, and strings. Explain what makes each process reversible or irreversible.

Q20.3 Household refrigerators have arrays or coils of tubing on the outside, usually at the back or bottom. When the refrigerator is running, the tubing becomes quite hot. Where does the heat come from?

Q20.4 Suppose you try to cool the kitchen of your house by leaving the refrigerator door open. What happens? Why? Would the result be the same if you left open a picnic cooler full of ice? Explain the reason for any differences.

Q20.5 Why must a room air conditioner be placed in a window rather than just set on the floor and plugged in? Why can a refrigerator be set on the floor and plugged in?

Q20.6 Is it a violation of the second law of thermodynamics to convert mechanical energy completely into heat? To convert heat completely into work? Explain your answers.

Q20.7 Imagine a special air filter placed in a window of a house. The tiny holes in the filter allow only air molecules moving faster than a certain speed to exit the house, and allow only air molecules moving slower than that speed to enter the house from outside. Explain why such an air filter would cool the house, and why the second law of thermodynamics makes building such a filter an impossible task.

Q20.8 An electric motor has its shaft coupled to that of an electric generator. The motor drives the generator, and some current from the generator is used to run the motor. The excess current is used to light a home. What is wrong with this scheme?

Q20.9 When a wet cloth is hung up in a hot wind in the desert, it is cooled by evaporation to a temperature that may be 20 C° or so below that of the air. Discuss this process in light of the second law of thermodynamics.

Q20.10 Compare the pV-diagram for the Otto cycle in Fig. 20.6 with the diagram for the Carnot heat engine in Fig. 20.13. Explain some of the important differences between the two cycles.

Q20.11 The efficiency of heat engines is high when the temperature difference between the hot and cold reservoirs is large. Refrigerators, on the other hand, work better when the temperature difference is small. Thinking of the mechanical refrigeration cycle shown in Fig. 20.9, explain in physical terms why it takes less work to remove heat from the working substance if the two reservoirs (the inside of the refrigerator and the outside air) are at nearly the same temperature, than if the outside air is much warmer than the interior of the refrigerator.

Q20.12 What would be the efficiency of a Carnot engine operating with $T_H = T_C$? What would be the efficiency if $T_C = 0$ K and T_H were any temperature above 0 K? Interpret your answers.

Q20.13 Real heat engines, like the gasoline engine in a car, always have some friction between their moving parts, although lubricants keep the friction to a minimum. Would a heat engine with completely frictionless parts be 100% efficient? Why or why not? Does the answer depend on whether or not the engine runs on the Carnot cycle? Again, why or why not?

Q20.14 Does a refrigerator full of food consume more power if the room temperature is 20°C than if it is 15°C? Or is the power consumption the same? Explain your reasoning.

Q20.15 In Example 20.4, a Carnot refrigerator requires a work input of only 230 J to extract 346 J of heat from the cold reservoir. Doesn't this discrepancy imply a violation of the law of conservation of energy? Explain why or why not.

Q20.16 How can the thermal conduction of heat from a hot object to a cold object increase entropy when the same amount of heat that flows out of the hot object flows into the cold one?

Q20.17 Explain why each of the following processes is an example of increasing randomness: mixing hot and cold water; free expansion of a gas; irreversible heat flow; developing heat by mechanical friction. Are entropy increases involved in all of these? Why or why not?

Q20.18 The free expansion of an ideal gas is an adiabatic process and so no heat is transferred. No work is done, so the internal energy does not change. Thus, $Q/T = 0$, yet the randomness of the system and thus its entropy have increased after the expansion. Why does Eq. (20.19) not apply to this situation?

Q20.19 Are the earth and sun in thermal equilibrium? Are there entropy changes associated with the transmission of energy from the sun to the earth? Does radiation differ from other modes of heat transfer with respect to entropy changes? Explain your reasoning.

Q20.20 Suppose that you put a hot object in thermal contact with a cold object and observe (much to your surprise) that heat flows from the cold object to the hot object, making the cold one colder and the hot one hotter. Does this process necessarily violate the first law of thermodynamics? The second law of thermodynamics? Explain.

Q20.21 If you run a movie film backward, it is as if the direction of time were reversed. In the time-reversed movie, would you see processes that violate conservation of energy? Conservation of linear momentum? Would you see processes that violate the second law of thermodynamics? In each case, if law-breaking processes could occur, give some examples.

Q20.22 BIO Some critics of biological evolution claim that it violates the second law of thermodynamics, since evolution involves simple life forms developing into more complex and more highly ordered organisms. Explain why this is not a valid argument against evolution.

Q20.23 BIO A growing plant creates a highly complex and organized structure out of simple materials such as air, water, and trace minerals. Does this violate the second law of thermodynamics? Why or why not? What is the plant's ultimate source of energy? Explain.

EXERCISES

Section 20.2 Heat Engines

20.1 • A diesel engine performs 2200 J of mechanical work and discards 4300 J of heat each cycle. (a) How much heat must be supplied to the engine in each cycle? (b) What is the thermal efficiency of the engine?

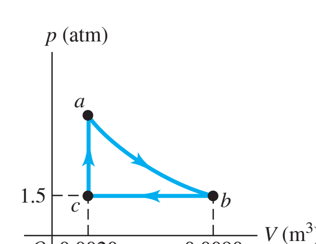
20.2 • An aircraft engine takes in 9000 J of heat and discards 6400 J each cycle. (a) What is the mechanical work output of the engine during one cycle? (b) What is the thermal efficiency of the engine?

20.3 • **A Gasoline Engine.** A gasoline engine takes in 1.61×10^4 J of heat and delivers 3700 J of work per cycle. The heat is obtained by burning gasoline with a heat of combustion of 4.60×10^4 J/g. (a) What is the thermal efficiency? (b) How much heat is discarded in each cycle? (c) What mass of fuel is burned in each cycle? (d) If the engine goes through 60.0 cycles per second, what is its power output in kilowatts? In horsepower?

20.4 • A gasoline engine has a power output of 180 kW (about 241 hp). Its thermal efficiency is 28.0%. (a) How much heat must be supplied to the engine per second? (b) How much heat is discarded by the engine per second?

20.5 •• The pV-diagram in **Fig. E20.5** shows a cycle of a heat engine that uses 0.250 mol of an ideal gas with $\gamma = 1.40$. Process ab is adiabatic. (a) Find the pressure of the gas at point a. (b) How much heat enters this gas per cycle, and where does it happen? (c) How much heat leaves this gas in a cycle, and where does it occur? (d) How much work does this engine do in a cycle? (e) What is the thermal efficiency of the engine?

Figure **E20.5**



20.6 •• **CP** A heat engine uses a large insulated tank of ice water as its cold reservoir. In 100 cycles the engine takes in 8000 J of heat energy from the hot reservoir and the rejected heat melts 0.0180 kg of ice in the tank. During these 100 cycles, how much work is performed by the engine?

Section 20.3 Internal-Combustion Engines

20.7 • The theoretical efficiency of an Otto cycle is 63%. The ratio of the C_p and C_V heat capacities for the gas in the engine is $\frac{7}{5}$. (a) What is the compression ratio r? (b) If the net work done per cycle is 12.6 kJ, how much heat is rejected and how much is absorbed by the engine during each cycle?

20.8 • (a) Calculate the theoretical efficiency for an Otto-cycle engine with $\gamma = 1.40$ and $r = 9.50$. (b) If this engine takes in 10,000 J of heat from burning its fuel, how much heat does it discard to the outside air?

20.9 •• The Otto-cycle engine in a Mercedes-Benz SLK230 has a compression ratio of 8.8. (a) What is the ideal efficiency of the engine? Use $\gamma = 1.40$. (b) The engine in a Dodge Viper GT2 has a slightly higher compression ratio of 9.6. How much increase in the ideal efficiency results from this increase in the compression ratio?

Section 20.4 Refrigerators

20.10 •• A freezer has a coefficient of performance of 2.40. The freezer is to convert 1.80 kg of water at 25.0°C to 1.80 kg of ice at −5.0°C in one hour. (a) What amount of heat must be removed from the water at 25.0°C to convert it to ice at −5.0°C? (b) How much electrical energy is consumed by the freezer during this hour? (c) How much wasted heat is delivered to the room in which the freezer sits?

20.11 •• A refrigerator has a coefficient of performance of 2.25, runs on an input of 135 W of electrical power, and keeps its inside compartment at 5°C. If you put a dozen 1.0 L plastic bottles of water at 31°C into this refrigerator, how long will it take for them to be cooled down to 5°C? (Ignore any heat that leaves the plastic.)

20.12 • A refrigerator has a coefficient of performance of 2.10. In each cycle it absorbs 3.10×10^4 J of heat from the cold reservoir. (a) How much mechanical energy is required each cycle to operate the refrigerator? (b) During each cycle, how much heat is discarded to the high-temperature reservoir?

Section 20.6 The Carnot Cycle

20.13 •• In a Carnot engine the hot reservoir is 72.0 C° warmer than the cold reservoir. The engine's efficiency is 12.5%. What are the Kelvin temperatures of the two reservoirs?

20.14 • A Carnot engine is operated between two heat reservoirs at temperatures of 520 K and 300 K. (a) If the engine receives 6.45 kJ of heat energy from the reservoir at 520 K in each cycle, how many joules per cycle does it discard to the reservoir at 300 K? (b) How much mechanical work is performed by the engine during each cycle? (c) What is the thermal efficiency of the engine?

20.15 • A Carnot engine whose high-temperature reservoir is at 620 K takes in 550 J of heat at this temperature in each cycle and gives up 335 J to the low-temperature reservoir. (a) How much mechanical work does the engine perform during each cycle? What is (b) the temperature of the low-temperature reservoir; (c) the thermal efficiency of the cycle?

20.16 •• An ice-making machine operates in a Carnot cycle. It takes heat from water at 0.0°C and rejects heat to a room at 24.0°C. Suppose that 85.0 kg of water at 0.0°C are converted to ice at 0.0°C. (a) How much heat is discharged into the room? (b) How much energy must be supplied to the device?

20.17 • A Carnot engine performs 2.5×10^4 J of work in each cycle and has an efficiency of 66%. (a) How much heat does the engine extract from its heat source in each cycle? (b) If the engine exhausts heat at room temperature (20.0°C), what is the temperature of its heat source?

20.18 •• A certain brand of freezer is advertised to use 730 kW·h of energy per year. (a) Assuming the freezer operates for 5 hours each day, how much power does it require while operating? (b) If the freezer keeps its interior at −5.0°C in a 20.0°C room, what is its theoretical maximum performance coefficient? (c) What is the theoretical maximum amount of ice this freezer could make in an hour, starting with water at 20.0°C?

20.19 • A Carnot refrigerator is operated between two heat reservoirs at temperatures of 320 K and 270 K. (a) If in each cycle the refrigerator receives 415 J of heat energy from the reservoir at 270 K, how many joules of heat energy does it deliver to the reservoir at 320 K? (b) If the refrigerator completes 165 cycles each minute, what power input is required to operate it? (c) What is the coefficient of performance of the refrigerator?

20.20 •• A Carnot heat engine uses a hot reservoir consisting of a large amount of boiling water and a cold reservoir consisting of a large tub of ice and water. In 5 minutes of operation, the heat rejected by the engine melts 0.0400 kg of ice. During this time, how much work W is performed by the engine?

20.21 •• Carnot refrigerator A has a 16% higher coefficient of performance than Carnot refrigerator B. The temperature difference between the hot and cold reservoirs is 30% greater for B than for A. If the cold-reservoir temperature for refrigerator B is 180 K, what is the cold-reservoir temperature for refrigerator A?

Section 20.7 Entropy

20.22 • **CALC** You decide to take a nice hot bath but discover that your thoughtless roommate has used up most of the hot water. You fill the tub with 195 kg of 30.0°C water and attempt to warm it further by pouring in 5.00 kg of boiling water from the stove. (a) Is this a reversible or an irreversible process? Use physical reasoning to explain. (b) Calculate the final temperature of the bath water. (c) Calculate the net change in entropy of the system (bath water + boiling water), assuming no heat exchange with the air or the tub itself.

20.23 • A sophomore with nothing better to do adds heat to 0.350 kg of ice at 0.0°C until it is all melted. (a) What is the change in entropy of the water? (b) The source of heat is a very massive object at 25.0°C. What is the change in entropy of this object? (c) What is the total change in entropy of the water and the heat source?

20.24 •• **CALC** You make tea with 0.250 kg of 85.0°C water and let it cool to room temperature (20.0°C). (a) Calculate the entropy change of the water while it cools. (b) The cooling process is essentially isothermal for the air in your kitchen. Calculate the change in entropy of the air while the tea cools, assuming that all of the heat lost by the water goes into the air. What is the total entropy change of the system tea + air?

20.25 • Three moles of an ideal gas undergo a reversible isothermal compression at 20.0°C. During this compression, 1850 J of work is done on the gas. What is the change of entropy of the gas?

20.26 •• What is the change in entropy of 0.130 kg of helium gas at the normal boiling point of helium when it all condenses isothermally to 1.00 L of liquid helium? (*Hint:* See Table 17.4 in Section 17.6.)

20.27 • (a) Calculate the change in entropy when 1.00 kg of water at 100°C is vaporized and converted to steam at 100°C (see Table 17.4). (b) Compare your answer to the change in entropy when 1.00 kg of ice is melted at 0°C, calculated in Example 20.5 (Section 20.7). Is the change in entropy greater for melting or for vaporization? Interpret your answer using the idea that entropy is a measure of the randomness of a system.

Section 20.8 Microscopic Interpretation of Entropy

20.28 • **CALC** Two moles of an ideal gas occupy a volume V. The gas expands isothermally and reversibly to a volume $3V$. (a) Is the velocity distribution changed by the isothermal expansion? Explain. (b) Use Eq. (20.23) to calculate the change in entropy of the gas. (c) Use Eq. (20.18) to calculate the change in entropy of the gas. Compare this result to that obtained in part (b).

20.29 • **CALC** A lonely party balloon with a volume of 2.40 L and containing 0.100 mol of air is left behind to drift in the temporarily uninhabited and depressurized International Space Station. Sunlight coming through a porthole heats and explodes the balloon, causing the air in it to undergo a free expansion into the empty station, whose total volume is 425 m^3. Calculate the entropy change of the air during the expansion.

PROBLEMS

20.30 • **CP** As a budding mechanical engineer, you are called upon to design a Carnot engine that has 2.00 mol of a monatomic ideal gas as its working substance and operates from a high-temperature reservoir at 500°C. The engine is to lift a 15.0 kg weight 2.00 m per cycle, using 500 J of heat input. The gas in the engine chamber can have a minimum volume of 5.00 L during the cycle. (a) Draw a pV-diagram for this cycle. Show in your diagram where heat enters and leaves the gas. (b) What must be the temperature of the cold reservoir? (c) What is the thermal efficiency of the engine? (d) How much heat energy does this engine waste per cycle? (e) What is the maximum pressure that the gas chamber will have to withstand?

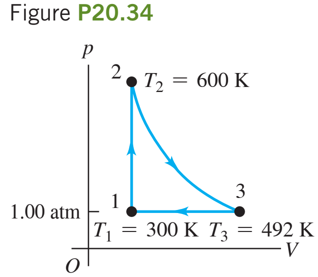
20.31 •• **CP** An ideal Carnot engine operates between 500°C and 100°C with a heat input of 250 J per cycle. (a) How much heat is delivered to the cold reservoir in each cycle? (b) What minimum number of cycles is necessary for the engine to lift a 500 kg rock through a height of 100 m?

20.32 •• **BIO** **Entropy of Metabolism.** An average sleeping person metabolizes at a rate of about 80 W by digesting food or burning fat. Typically, 20% of this energy goes into bodily functions, such as cell repair, pumping blood, and other uses of mechanical energy, while the rest goes to heat. Most people get rid of all this excess heat by transferring it (by conduction and the flow of blood) to the surface of the body, where it is radiated away. The normal internal temperature of the body (where the metabolism takes place) is 37°C, and the skin is typically 7 C° cooler. By how much does the person's entropy change per second due to this heat transfer?

20.33 •• **CP** A certain heat engine operating on a Carnot cycle absorbs 410 J of heat per cycle at its hot reservoir at 135°C and has a thermal efficiency of 22.0%. (a) How much work does this engine do per cycle? (b) How much heat does the engine waste each cycle? (c) What is the temperature of the cold reservoir? (d) By how much does the engine change the entropy of the world each cycle? (e) What mass of water could this engine pump per cycle from a well 35.0 m deep?

20.34 • A heat engine takes 0.350 mol of a diatomic ideal gas around the cycle shown in the pV-diagram of **Fig. P20.34**. Process $1 \rightarrow 2$ is at constant volume, process $2 \rightarrow 3$ is adiabatic, and process $3 \rightarrow 1$ is at a constant pressure of 1.00 atm. The value of γ for this gas is 1.40. (a) Find the pressure and volume at points 1, 2, and 3. (b) Calculate Q, W, and ΔU for each of the three processes. (c) Find the net work done by the gas in the cycle. (d) Find the net heat flow into the engine in one cycle. (e) What is the thermal efficiency of the engine? How does this compare to the efficiency of a Carnot-cycle engine operating between the same minimum and maximum temperatures T_1 and T_2?

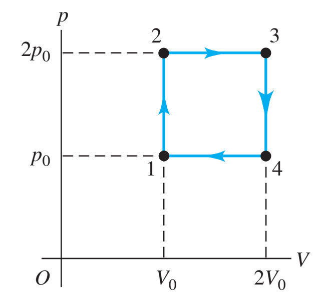
Figure P20.34



20.35 •• **BIO** **Entropy Change from Digesting Fat.** Digesting fat produces 9.3 food calories per gram of fat, and typically 80% of this energy goes to heat when metabolized. (One food calorie is 1000 calories and therefore equals 4186 J.) The body then moves all this heat to the surface by a combination of thermal conductivity and motion of the blood. The internal temperature of the body (where digestion occurs) is normally 37°C, and the surface is usually about 30°C. By how much do the digestion and metabolism of a 2.50 g pat of butter change your body's entropy? Does it increase or decrease?

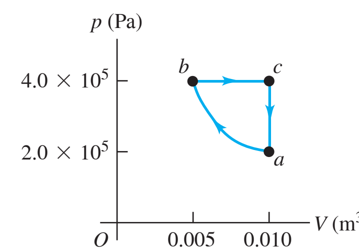
20.36 •• What is the thermal efficiency of an engine that takes n moles of diatomic ideal gas through the cycle $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$ shown in **Fig. P20.36**?

Figure P20.36



20.37 •• **CALC** You build a heat engine that takes 1.00 mol of an ideal diatomic gas through the cycle shown in **Fig. P20.37**. (a) Show that process ab is an isothermal compression. (b) During which process(es) of the cycle is heat absorbed by the gas? During which process(es) is heat rejected? How do you know? Calculate (c) the temperature at points a, b, and c; (d) the net heat exchanged with the surroundings and net work done by the engine in one cycle; (e) the thermal efficiency of the engine.

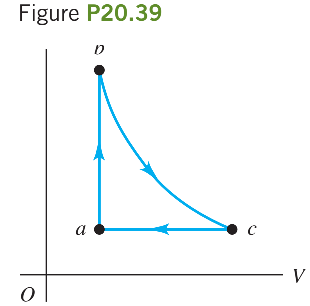
Figure P20.37



20.38 •• **CALC** One winter evening you fill a pint-sized plastic cup with ice and water to drink before going to bed. Estimate the volume of the ice and the volume of the water in the cup. Use the densities of ice and liquid water to find the mass of each. You set the cup on the kitchen counter, forget about it, and go to bed. The air in the kitchen is maintained at a constant 22.2°C by the central heating system of your apartment. When you wake up in the morning, the cup contains liquid water at 22.2°C. (a) What is the entropy change of the ice water during the process in which it went from 0°C to liquid water at 22.2°C as you slept? (b) What is the entropy change of the air in the room due to the heat that flows from it into the contents of the cup? (c) What is the total entropy change of the air–ice water system due to the heat exchange that changes the phase and temperature of the ice water? Is the total entropy change positive or negative?

20.39 • **CALC** A heat engine operates using the cycle shown in **Fig. P20.39**. The working substance is 2.00 mol of helium gas, which reaches a maximum temperature of 327°C. Assume the helium can be treated as an ideal gas. Process bc is isothermal. The pressure in states a and c is 1.00×10^5 Pa, and the pressure in state b is 3.00×10^5 Pa. (a) How much heat enters the gas and how much leaves the gas each cycle? (b) How much work does the engine do each cycle, and what is its efficiency? (c) Compare this engine's efficiency with the maximum possible efficiency attainable with the hot and cold reservoirs used by this cycle.

Figure P20.39

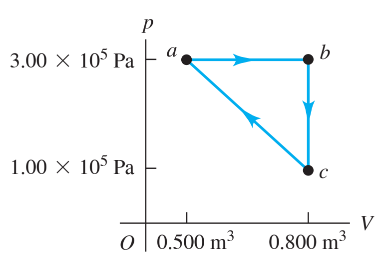


20.40 •• CP BIO A Human Engine. You decide to use your body as a Carnot heat engine. The operating gas is in a tube with one end in your mouth (where the temperature is 37.0°C) and the other end at the surface of your skin, at 30.0°C. (a) What is the maximum efficiency of such a heat engine? Would it be a very useful engine? (b) Suppose you want to use this human engine to lift a 2.50 kg box from the floor to a tabletop 1.20 m above the floor. How much must you increase the gravitational potential energy, and how much heat input is needed to accomplish this? (c) If your favorite candy bar has 350 food calories (1 food calorie = 4186 J) and 80% of the food energy goes into heat, how many of these candy bars must you eat to lift the box in this way?

20.41 ••• An experimental power plant at the Natural Energy Laboratory of Hawaii generates electricity from the temperature gradient of the ocean. The surface and deep-water temperatures are 27°C and 6°C, respectively. (a) What is the maximum theoretical efficiency of this power plant? (b) If the power plant is to produce 210 kW of power, at what rate must heat be extracted from the warm water? At what rate must heat be absorbed by the cold water? Assume the maximum theoretical efficiency. (c) The cold water that enters the plant leaves it at a temperature of 10°C. What must be the flow rate of cold water through the system? Give your answer in kg/h and in L/h.

20.42 • A monatomic ideal gas is taken around the cycle shown in **Fig. P20.42** in the direction shown in the figure. The path for process $c \rightarrow a$ is a straight line in the pV-diagram. (a) Calculate Q, W, and ΔU for each process $a \rightarrow b$, $b \rightarrow c$, and $c \rightarrow a$. (b) What are Q, W, and ΔU for one complete cycle? (c) What is the efficiency of the cycle?

Figure **P20.42**



20.43 • CALC A cylinder contains oxygen at a pressure of 2.00 atm. The volume is 4.00 L, and the temperature is 300 K. Assume that the oxygen may be treated as an ideal gas. The oxygen is carried through the following processes:

(i) Heated at constant pressure from the initial state (state 1) to state 2, which has $T = 450$ K.
(ii) Cooled at constant volume to 250 K (state 3).
(iii) Compressed at constant temperature to a volume of 4.00 L (state 4).
(iv) Heated at constant volume to 300 K, which takes the system back to state 1.

(a) Show these four processes in a pV-diagram, giving the numerical values of p and V in each of the four states. (b) Calculate Q and W for each of the four processes. (c) Calculate the net work done by the oxygen in the complete cycle. (d) What is the efficiency of this device as a heat engine? How does this compare to the efficiency of a Carnot-cycle engine operating between the same minimum and maximum temperatures of 250 K and 450 K?

20.44 ••• A typical coal-fired power plant generates 1000 MW of usable power at an overall thermal efficiency of 40%. (a) What is the rate of heat input to the plant? (b) The plant burns anthracite coal, which has a heat of combustion of 2.65×10^7 J/kg. How much coal does the plant use per day, if it operates continuously? (c) At what rate is heat ejected into the cool reservoir, which is the nearby river? (d) The river is at 18.0°C before it reaches the power plant and 18.5°C after it has received the plant's waste heat. Calculate the river's flow rate, in cubic meters per second. (e) By how much does the river's entropy increase each second?

20.45 •• A Carnot engine operates between two heat reservoirs at temperatures T_H and T_C. An inventor proposes to increase the efficiency by running one engine between T_H and an intermediate temperature T' and a second engine between T' and T_C, using as input the heat expelled by the first engine. Compute the efficiency of this composite system, and compare it to that of the original engine.

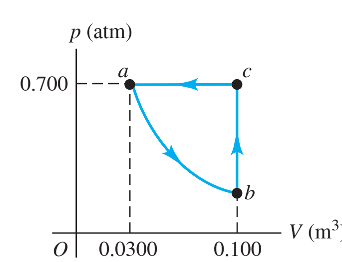
20.46 •• (a) Estimate the fuel efficiency, in miles per gallon, of a typical car. (b) The combustion of a gallon of gasoline releases 120 MJ of energy. If a typical car is driven at 40 mph and the engine is turning at 3000 revolutions per minute, then based on your estimate of fuel efficiency, how many joules of energy are released by combustion during each cycle of the engine? (c) A typical car engine runs with 20% efficiency. Based on this figure, how many watts of power does the car supply when moving at a steady 40 mph? (d) What is that figure in horsepower?

20.47 • Automotive Thermodynamics. A Volkswagen Passat has a six-cylinder Otto-cycle engine with compression ratio $r = 10.6$. The diameter of each cylinder, called the *bore* of the engine, is 82.5 mm. The distance that the piston moves during the compression in Fig. 20.5, called the *stroke* of the engine, is 86.4 mm. The initial pressure of the air–fuel mixture (at point a in Fig. 20.6) is 8.50×10^4 Pa, and the initial temperature is 300 K (the same as the outside air). Assume that 200 J of heat is added to each cylinder in each cycle by the burning gasoline, and that the gas has $C_V = 20.5$ J/mol·K and $\gamma = 1.40$. (a) Calculate the total work done in one cycle in each cylinder of the engine, and the heat released when the gas is cooled to the temperature of the outside air. (b) Calculate the volume of the air–fuel mixture at point a in the cycle. (c) Calculate the pressure, volume, and temperature of the gas at points b, c, and d in the cycle. In a pV-diagram, show the numerical values of p, V, and T for each of the four states. (d) Compare the efficiency of this engine with the efficiency of a Carnot-cycle engine operating between the same maximum and minimum temperatures.

20.48 •• (a) Estimate the size of your room. (b) One mole of an ideal gas at STP occupies a volume of 22.4 L. Using this figure, estimate the number of moles of air in your room. (c) How many molecules is that? (d) There are N^N ways to distribute N things onto N sites. As a rough approximation, assume there are as many states for air molecules as there are air molecules in your room. Compute the entropy of the air in your room.

20.49 •• The pV-diagram in **Fig. P20.49** shows the cycle for a refrigerator operating on 0.850 mol of H_2. Assume that the gas can be treated as ideal. Process ab is isothermal. Find the coefficient of performance of this refrigerator.

Figure **P20.49**



20.50 •• BIO Human Entropy. A person with skin of surface area 1.85 m^2 and temperature 30.0°C is resting in an insulated room where the ambient air temperature is 20.0°C. Assume that this person gets rid of excess heat by radiation only. By how much does the person change the entropy of the air in this room each second? (Recall that the room radiates back into the person and that the emissivity of the skin is 1.00.)

20.51 •• CALC An object of mass m_1, specific heat c_1, and temperature T_1 is placed in contact with a second object of mass m_2, specific heat c_2, and temperature $T_2 > T_1$. As a result, the temperature of the first object increases to T and the temperature of the second object decreases to T'. (a) Show that the entropy increase of the system is

$$\Delta S = m_1 c_1 \ln \frac{T}{T_1} + m_2 c_2 \ln \frac{T'}{T_2}$$

and show that energy conservation requires that

$$m_1 c_1 (T - T_1) = m_2 c_2 (T_2 - T')$$

(b) Show that the entropy change ΔS, considered as a function of T, is a *maximum* if $T = T'$, which is just the condition of thermodynamic equilibrium. (c) Discuss the result of part (b) in terms of the idea of entropy as a measure of randomness.

20.52 •• CALC To heat 1 cup of water $(250\ \text{cm}^3)$ to make coffee, you place an electric heating element in the cup. As the water temperature increases from 20°C to 78°C, the temperature of the heating element remains at a constant 120°C. Calculate the change in entropy of (a) the water; (b) the heating element; (c) the system of water and heating element. (Make the same assumption about the specific heat of water as in Example 20.10 in Section 20.7, and ignore the heat that flows into the ceramic coffee cup itself.) (d) Is this process reversible or irreversible? Explain.

20.53 •• DATA In your summer job with a venture capital firm, you are given funding requests from four inventors of heat engines. The inventors claim the following data for their operating prototypes:

	Prototype			
	A	**B**	**C**	**D**
T_C (°C), low-temperature reservoir	47	17	−33	37
T_H (°C), high-temperature reservoir	177	197	247	137
Claimed efficiency e (%)	21	35	56	20

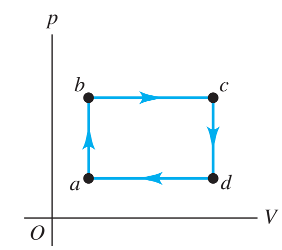
(a) Based on the T_C and T_H values for each prototype, find the maximum possible efficiency for each. (b) Are any of the claimed efficiencies impossible? Explain. (c) For all prototypes with an efficiency that is possible, rank the prototypes in decreasing order of the ratio of claimed efficiency to maximum possible efficiency.

20.54 •• DATA For a refrigerator or air conditioner, the coefficient of performance K (often denoted as COP) is, as in Eq. (20.9), the ratio of cooling output $|Q_C|$ to the required electrical energy input $|W|$, both in joules. The coefficient of performance is also expressed as a ratio of powers,

$$K = \frac{|Q_C|/t}{|W|/t}$$

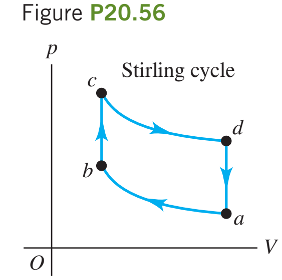
where $|Q_C|/t$ is the cooling power and $|W|/t$ is the electrical power input to the device, both in watts. The energy efficiency ratio (EER) is the same quantity expressed in units of Btu for $|Q_C|$ and W • h for $|W|$. (a) Derive a general relationship that expresses EER in terms of K. (b) For a home air conditioner, EER is generally determined for a 95°F outside temperature and an 80°F return air temperature. Calculate EER for a Carnot device that operates between 95°F and 80°F. (c) You have an air conditioner with an EER of 10.9. Your home on average requires a total cooling output of $|Q_C| = 1.9 \times 10^{10}$ J per year. If electricity costs you 15.3 cents per kW • h, how much do you spend per year, on average, to operate your air conditioner? (Assume that the unit's EER accurately represents the operation of your air conditioner. A *seasonal energy efficiency ratio* (SEER) is often used. The SEER is calculated over a range of outside temperatures to get a more accurate seasonal average.) (d) You are considering replacing your air conditioner with a more efficient one with an EER of 14.6. Based on the EER, how much would that save you on electricity costs in an average year?

Figure **P20.55**



20.55 ••• DATA You are conducting experiments to study prototype heat engines. In one test, 4.00 mol of argon gas are taken around the cycle shown in **Fig. P20.55**. The pressure is low enough for the gas to be treated as ideal. You measure the gas temperature in states a, b, c, and d and find $T_a = 250.0$ K, $T_b = 300.0$ K, $T_c = 380.0$ K, and $T_d = 316.7$ K. (a) Calculate the efficiency e of the cycle. (b) Disappointed by the cycle's low efficiency, you consider doubling the number of moles of gas while keeping the pressure and volume the same. What would e be then? (c) You remember that the efficiency of a Carnot cycle increases if the temperature of the hot reservoir is increased. So, you return to using 4.00 mol of gas but double the volume in states c and d while keeping the pressures the same. The resulting temperatures in these states are $T_c = 760.0$ K and $T_d = 633.4$ K. T_a and T_b remain the same as in part (a). Calculate e for this cycle with the new T_c and T_d values. (d) Encouraged by the increase in efficiency, you raise T_c and T_d still further. But e doesn't increase very much; it seems to be approaching a limiting value. If $T_a = 250.0$ K and $T_b = 300.0$ K and you keep volumes V_a and V_b the same as in part (a), then $T_c/T_d = T_b/T_a$ and $T_c = 1.20T_d$. Derive an expression for e as a function of T_d for this cycle. What value does e approach as T_d becomes very large?

Figure **P20.56**



20.56 •• CP In an ideal Stirling engine, n moles of an ideal gas are isothermally compressed, pressurized at constant volume, allowed to expand isothermally, and then cooled at constant volume, as shown in **Fig. P20.56**. The expansion takes place at temperature T_H and the compression at temperature T_C. (a) Show that $p_c/p_d = p_b/p_a$. This quantity is called the compression ratio of the engine, or the CR. Determine the work done in terms of the CR (b) during the expansion phase and (c) during the compression phase. (d) An ideal Stirling engine uses 1 mole of helium as its working substance and has a CR of 10, temperature reservoirs at 100ºC and 20ºC, and a frequency of operation of 100 Hz. What is its power output?

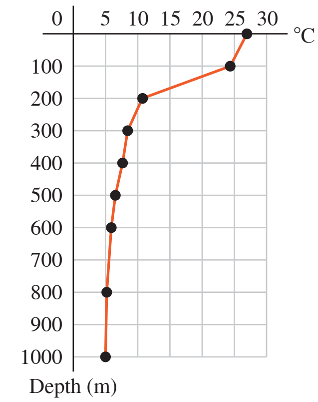
CHALLENGE PROBLEMS

20.57 ••• On planet Tralfamadore the cool atmosphere is rich in froklom molecules. Spherical life forms called sporons feed on these molecules and obtain energy through a Carnot cycle. When a sporon that has injested froklom gas has isothermally contracted to a trigger volume, it violently squeezes itself, reducing its volume by half. At this point, its interior achieves the combustion temperature at which froklom fuses exothermically to produce klombactic acid. At the point of combustion, a sporon has radius 8.0 cm and internal pressure 20.3 kPa. The sporon swells during the reaction period, maintaining constant temperature until its injested froklom is depleted. The sporon continues to swell rapidly until its internal temperature matches the ambient temperature of 123 K and its gauge pressure is zero. Then it contracts slowly, emitting klombactic acid while simultaneously injesting froklom from the atmosphere. When its radius shrinks by a factor of 3 it reaches the trigger volume, which starts another cycle. Aside from the sporon's skin membrane, all substances behave like ideal gases. Froklom gas has a ratio of heat capacities of 1.30. (a) What is the combustion temperature at which froklom fuses? (b) What is the ambient pressure on Tralfamadore? (c) If a sporon generates 60 kJ of work each hour, how much heat does it draw in per minute? (d) How much heat is rejected back to the atmosphere per minute?

20.58 ••• Consider a Diesel cycle that starts (at point a in Fig. 20.7) with air at temperature T_a. The air may be treated as an ideal gas. (a) If the temperature at point c is T_c, derive an expression for the efficiency of the cycle in terms of the compression ratio r. (b) What is the efficiency if $T_a = 300$ K, $T_c = 950$ K, $\gamma = 1.40$, and $r = 21.0$?

MCAT-STYLE PASSAGE PROBLEMS

Power from the Sea. *Ocean thermal energy conversion* is a process that uses the temperature difference between the warm surface water of tropical oceans and the cold deep-ocean water to run a heat engine. The graph shows a typical decrease of temperature with depth below the surface in tropical oceans. In the heat engine, the warmer surface water vaporizes a low-boiling-point fluid, such as ammonia. The heat of vaporization of ammonia is 260 cal/g at 27°C, the surface-water temperature. The vapor is used to turn a turbine and is then condensed back into a liquid by means of cold water



brought from deep below the surface through a large intake pipe. A power plant producing 10 MW of useful power would require a cold seawater flow rate of about 30,000 kg/s.

20.59 If the power plant uses a Carnot cycle and the desired theoretical efficiency is 6.5%, from what depth must cold water be brought? (a) 100 m; (b) 400 m; (c) 800 m; (d) deeper than 1000 m.

20.60 What is the change in entropy of the ammonia vaporized per second in the 10 MW power plant, assuming an ideal Carnot efficiency of 6.5%? (a) $+6 \times 10^6$ J/K per second; (b) $+5 \times 10^5$ J/K per second; (c) $+1 \times 10^5$ J/K per second; (d) 0.

20.61 Compare the entropy change of the warmer water to that of the colder water during one cycle of the heat engine, assuming an ideal Carnot cycle. (a) The entropy does not change during one cycle in either case. (b) The entropy of both increases, but the entropy of the colder water increases by more because its initial temperature is lower. (c) The entropy of the warmer water decreases by more than the entropy of the colder water increases, because some of the heat removed from the warmer water goes to the work done by the engine. (d) The entropy of the warmer water decreases by the same amount that the entropy of the colder water increases.

20.62 If the proposed plant is built and produces 10 MW but the rate at which waste heat is exhausted to the cold water is 165 MW, what is the plant's actual efficiency? (a) 5.7%; (b) 6.1%; (c) 6.5%; (d) 16.5%.

ANSWERS

Chapter Opening Question ?

(i) This is what a refrigerator does: It makes heat flow from the cold interior of the refrigerator to the warm outside. The second law of thermodynamics says that heat cannot *spontaneously* flow from a cold object to a hot one. A refrigerator has a motor that does work on the system to *force* the heat to flow in that way.

Key Example ✓ARIATION Problems

VP20.1.1 (a) 6.89×10^4 J (b) 5.65×10^4 J
VP20.1.2 0.173
VP20.1.3 (a) 6.17×10^8 J (b) 1.21×10^8 J
VP20.1.4 (a) 3.96×10^8 J (b) 1.7×10^9 J (c) 0.23
VP20.4.1 (a) 0.210 (b) 6.32×10^4 J (c) 104°C
VP20.4.2 (a) 0.60
(b) $W_{ab} = 576$ J, $W_{bc} = 1250$ J, $W_{cd} = -231$ J, $W_{da} = -1250$ J
VP20.4.3 (a) 7.52 (b) 5.32×10^5 J
VP20.4.4 (a) $nRT_H \ln 2$ (b) $\frac{3}{2}nR(T_H - T_C)$
(c) $T_C/T_H = 1 - \frac{2}{3}\ln 2 = 0.538$, $e = \frac{2}{3}\ln 2 = 0.462$
VP20.10.1 (a) 6.55×10^2 J/K (b) 1.92×10^3 J/K (c) 2.43×10^3 J/K
VP20.10.2 (a) +45.7 J/K (b) +45.7 J/K
VP20.10.3 (a) zero (b) −150 J/K (c) +218 J/K (d) +68 J/K; increase
VP20.10.4 (a) $+1.22 \times 10^3$ J/K (b) -1.05×10^3 J/K
(c) $+1.7 \times 10^2$ J/K; increase

Bridging Problem

(a) 34.8°C
(b) +10 J/K

? Water makes life possible: The cells of your body could not function without water in which to dissolve essential biological molecules. Water is such a good solvent because its molecules (i) have zero net charge; (ii) have zero net charge, but the positive and negative charges are separated; (iii) have nonzero net charge; (iv) do not respond to electric forces; (v) exert repulsive electric forces on each other.

21 Electric Charge and Electric Field

LEARNING OUTCOMES

In this chapter, you'll learn...

21.1 The nature of electric charge, and how we know that electric charge is conserved.
21.2 How objects become electrically charged.
21.3 How to use Coulomb's law to calculate the electric force between charges.
21.4 The distinction between electric force and electric field.
21.5 How to calculate the electric field due to a collection of charges.
21.6 How to use the idea of electric field lines to visualize and interpret electric fields.
21.7 How to calculate the properties of electric dipoles.

You'll need to review...

1.7–1.10 Vector algebra, including the scalar (dot) product and the vector (cross) product.
4.3 Newton's second law.
7.5 Stable and unstable equilibria.
12.5 Streamlines in fluid flow.

In Chapter 5 we mentioned the four kinds of fundamental forces. To this point the only one of these forces that we have examined in any detail is gravity. Now we are ready to examine the force of *electromagnetism,* which encompasses both electricity and magnetism.

Electromagnetic interactions involve particles that have *electric charge,* an attribute that is as fundamental as mass. Just as objects with mass are accelerated by gravitational forces, so electrically charged objects are accelerated by electric forces. The shock you feel when you scuff your shoes across a carpet and then reach for a metal doorknob is due to charged particles leaping between your finger and the doorknob. Electric currents are simply streams of charged particles flowing within wires in response to electric forces. Even the forces that hold atoms together to form solid matter, and that keep the atoms of solid objects from passing through each other, are fundamentally due to electric interactions between the charged particles within atoms.

We begin our study of electromagnetism in this chapter by examining the nature of electric charge. We'll find that charge is quantized and obeys a conservation principle. When charges are at rest in our frame of reference, they exert *electrostatic* forces on each other. These forces are of tremendous importance in chemistry and biology and have many technological applications. Electrostatic forces are governed by a simple relationship known as *Coulomb's law* and are most conveniently described by using the concept of *electric field.* In later chapters we'll expand our discussion to include electric charges in motion. This will lead us to an understanding of magnetism and, remarkably, of the nature of light.

While the key ideas of electromagnetism are conceptually simple, applying them to practical problems will make use of many of your mathematical skills, especially your knowledge of geometry and integral calculus. For this reason you may find this chapter and those that follow to be more mathematically demanding than earlier chapters. The reward for your extra effort will be a deeper understanding of principles that are at the heart of modern physics and technology.

21.1 ELECTRIC CHARGE

The ancient Greeks discovered as early as 600 B.C. that after they rubbed amber with wool, the amber could attract other objects. Today we say that the amber has acquired a net **electric charge,** or has become *charged.* The word "electric" is derived from the Greek word *elektron,* meaning amber. When you scuff your shoes across a nylon carpet, you become electrically charged, and you can charge a comb by passing it through dry hair.

Plastic rods and fur (real or fake) are particularly good for demonstrating **electrostatics,** the interactions between electric charges that are at rest (or nearly so). After we charge both plastic rods in **Fig. 21.1a** by rubbing them with the piece of fur, we find that the rods repel each other.

When we rub glass rods with silk, the glass rods also become charged and repel each other (Fig. 21.1b). But a charged plastic rod *attracts* a charged glass rod; furthermore, the plastic rod and the fur attract each other, and the glass rod and the silk attract each other (Fig. 21.1c).

These experiments and many others like them have shown that there are exactly two kinds of electric charge: the kind on the plastic rod rubbed with fur and the kind on the glass rod rubbed with silk. Benjamin Franklin (1706–1790) suggested calling these two kinds of charge *negative* and *positive,* respectively, and these names are still used. The plastic rod and the silk have negative charge; the glass rod and the fur have positive charge.

Two positive charges or two negative charges repel each other. A positive charge and a negative charge attract each other.

CAUTION Electric attraction and repulsion The attraction and repulsion of two charged objects are sometimes summarized as "Like charges repel, and opposite charges attract." But "like charges" does *not* mean that the two charges are exactly identical, only that both charges have the same algebraic *sign* (both positive or both negative). "Opposite charges" means that both objects have an electric charge, and those charges have different signs (one positive and the other negative).

Figure 21.1 Experiments in electrostatics. (a) Negatively charged objects repel each other. (b) Positively charged objects repel each other. (c) Positively charged objects and negatively charged objects attract each other.

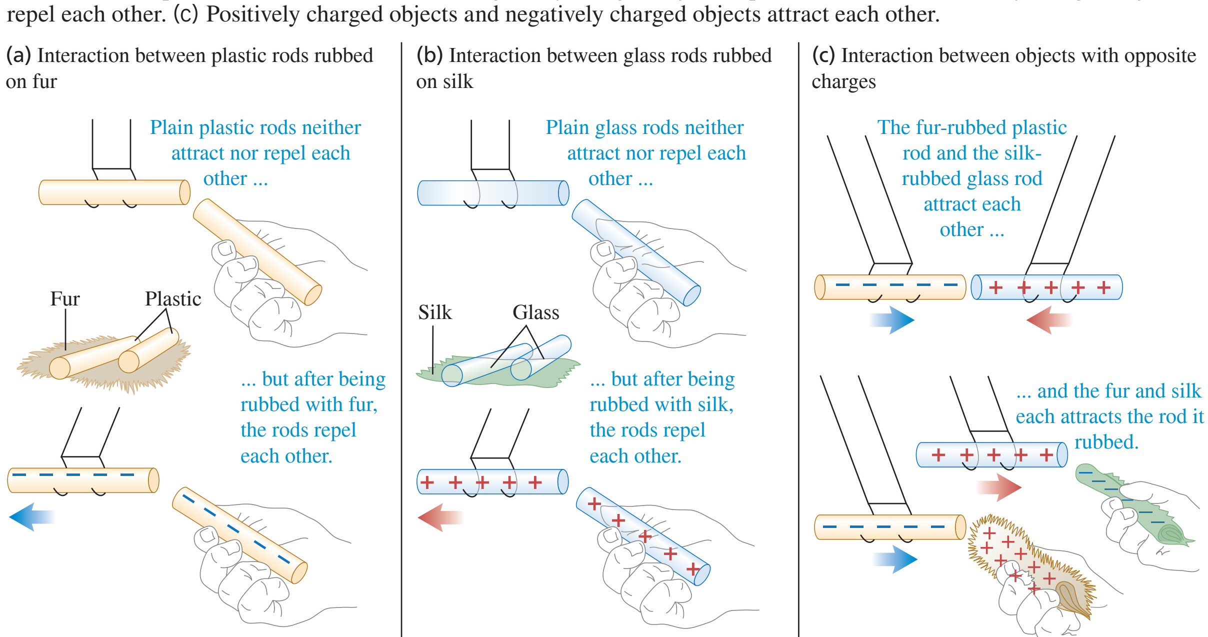
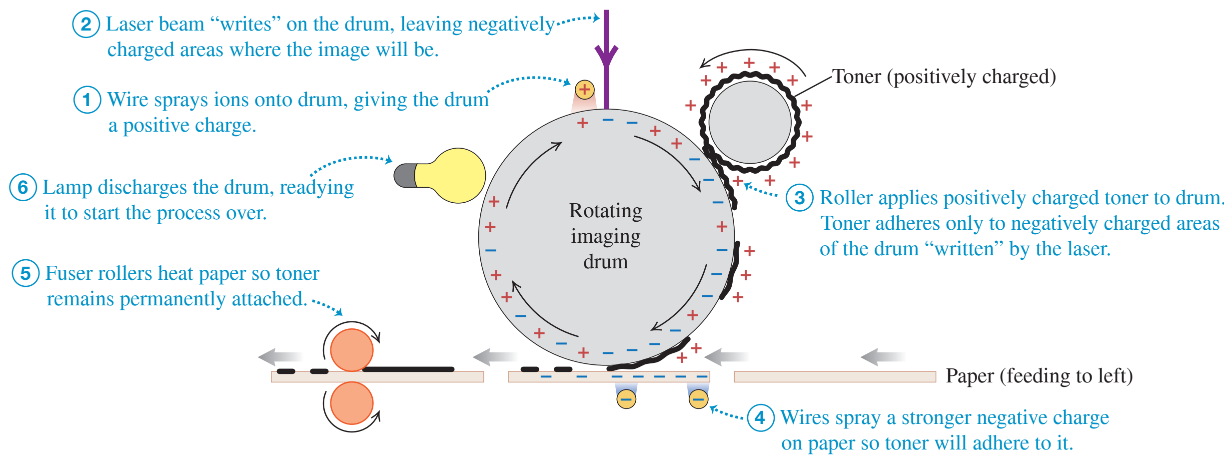


Figure **21.2** Schematic diagram of the operation of a laser printer.



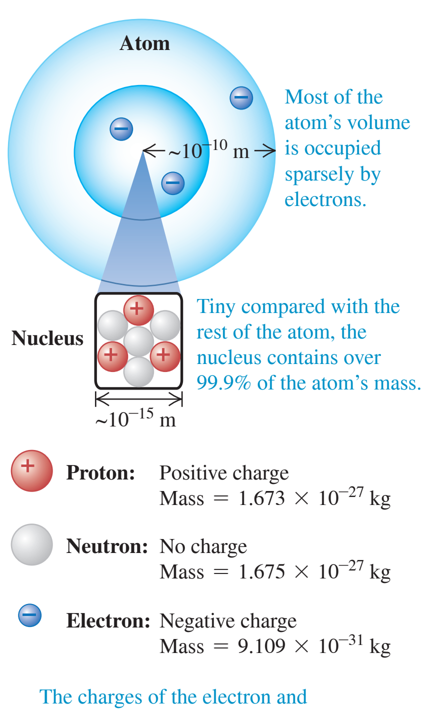
A laser printer (**Fig. 21.2**) utilizes the forces between charged objects. The printer's light-sensitive imaging drum is given a positive charge. As the drum rotates, a laser beam shines on selected areas of the drum, leaving those areas with a *negative* charge. Positively charged particles of toner adhere only to the areas of the drum "written" by the laser. When a piece of paper is placed in contact with the drum, the toner particles stick to the paper and form an image.

Electric Charge and the Structure of Matter

When you charge a rod by rubbing it with fur or silk as in Fig. 21.1, there is no visible change in the appearance of the rod. What, then, actually happens to the rod when you charge it? To answer this question, we must look more closely at the structure of atoms, the building blocks of ordinary matter.

The structure of atoms can be described in terms of three particles: the negatively charged **electron,** the positively charged **proton,** and the uncharged **neutron** (**Fig. 21.3**). The proton and neutron are combinations of other entities called *quarks,* which have charges of $\pm\frac{1}{3}$ and $\pm\frac{2}{3}$ times the electron charge. Isolated quarks have not been observed, and there are theoretical reasons to believe that it is impossible in principle to observe a quark in isolation.

Figure **21.3** The structure of an atom. The particular atom depicted here is lithium (see Fig. 21.4a).



The protons and neutrons in an atom make up a small, very dense core called the **nucleus,** with dimensions of the order of 10^{-15} m. Surrounding the nucleus are the electrons, extending out to distances of the order of 10^{-10} m from the nucleus. If an atom were a few kilometers across, its nucleus would be the size of a tennis ball. The negatively charged electrons are held within the atom by the attractive electric forces exerted on them by the positively charged nucleus. (The protons and neutrons are held within stable atomic nuclei by an attractive interaction, called the *strong nuclear force,* that overcomes the electric repulsion of the protons. The strong nuclear force has a short range, and its effects do not extend far beyond the nucleus.)

The masses of the individual particles, to the precision that they are known (as of this writing), are

$$\text{Mass of electron} = m_e = 9.10938356(11) \times 10^{-31}\ \text{kg}$$

$$\text{Mass of proton} = m_p = 1.672621898(21) \times 10^{-27}\ \text{kg}$$

$$\text{Mass of neutron} = m_n = 1.674927471(21) \times 10^{-27}\ \text{kg}$$

The numbers in parentheses are the uncertainties in the last two digits. Note that the masses of the proton and neutron are nearly equal and are roughly 2000 times the mass of the electron. Over 99.9% of the mass of any atom is concentrated in its nucleus.

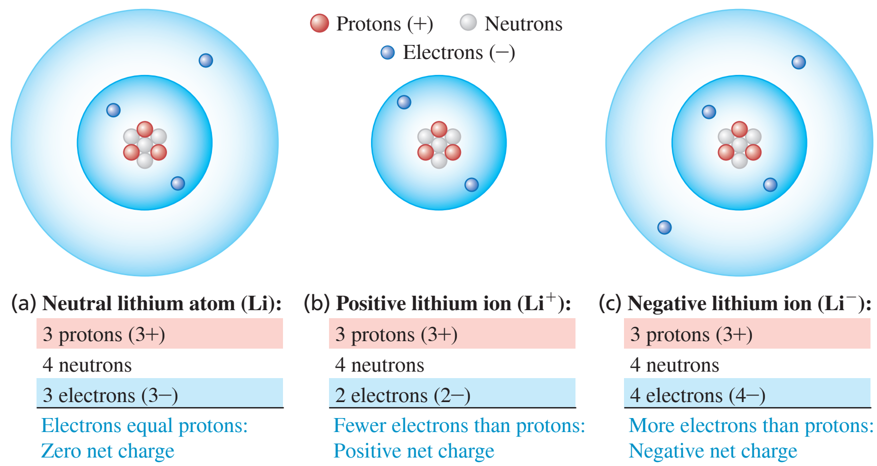


Figure **21.4** (a) A neutral atom has as many electrons as it does protons. (b) A positive ion has a deficit of electrons. (c) A negative ion has an excess of electrons. (The electron "shells" are a schematic representation of the actual electron distribution, a diffuse cloud many times larger than the nucleus.)

The negative charge of the electron has (within experimental error) *exactly* the same magnitude as the positive charge of the proton. In a neutral atom the number of electrons equals the number of protons in the nucleus, and the net electric charge (the algebraic sum of all the charges) is exactly zero (**Fig. 21.4a**). The number of protons or electrons in a neutral atom of an element is called the **atomic number** of the element. If one or more electrons are removed from an atom, what remains is called a **positive ion** (Fig. 21.4b). A **negative ion** is an atom that has *gained* one or more electrons (Fig. 21.4c). This gain or loss of electrons is called **ionization.**

When the total number of protons in a macroscopic object equals the total number of electrons, the total charge is zero and the object as a whole is electrically neutral. To give an object an excess negative charge, we may either *add negative* charges to a neutral object or *remove positive* charges from that object. Similarly, we can create an excess positive charge by either *adding positive* charge or *removing negative* charge. In most cases, negatively charged (and highly mobile) electrons are added or removed, and a "positively charged object" is one that has lost some of its normal complement of electrons. When we speak of the charge of an object, we always mean its *net* charge. The net charge is always a very small fraction (typically no more than 10^{-12}) of the total positive charge or negative charge in the object.

Electric Charge Is Conserved

Implicit in the foregoing discussion are two very important principles, the first of which is:

> **PRINCIPLE OF CONSERVATION OF CHARGE** **The algebraic sum of all the electric charges in any closed system is constant.**

If we rub together a plastic rod and a piece of fur, both initially uncharged, the rod acquires a negative charge (since it takes electrons from the fur) and the fur acquires a positive charge of the *same* magnitude (since it has lost as many electrons as the rod has gained). Hence the total electric charge on the two objects together does not change. In any charging process, charge is not created or destroyed; it is merely *transferred* from one object to another.

Conservation of charge is thought to be a *universal* conservation law. No experimental evidence for any violation of this principle has ever been observed. Even in high-energy interactions in which particles are created and destroyed, such as the creation of electron–positron pairs, the total charge of any closed system is exactly constant.

Figure 21.5 Most of the forces on this water skier are electric. Electric interactions between adjacent molecules give rise to the force of the water on the ski, the tension in the tow rope, and the resistance of the air on the skier's body. Electric interactions also hold the atoms of the skier's body together. Only one wholly nonelectric force acts on the skier: the force of gravity.

The second important principle is:

The magnitude of charge of the electron or proton is a natural unit of charge.

Every observable amount of electric charge is always an integer multiple of this basic unit. We say that charge is *quantized*. A familiar example of quantization is money. When you pay cash for an item in a store, you have to do it in one-cent increments. Cash can't be divided into amounts smaller than one cent, and electric charge can't be divided into amounts smaller than the charge of one electron or proton. (The quark charges, $\pm\frac{1}{3}$ and $\pm\frac{2}{3}$ of the electron charge, are probably not observable as isolated charges.) Thus the charge on any macroscopic object is always zero or an integer multiple (negative or positive) of the electron charge.

Understanding the electric nature of matter gives us insight into many aspects of the physical world (**Fig. 21.5**). The chemical bonds that hold atoms together to form molecules are due to electric interactions between the atoms. They include the strong ionic bonds that hold sodium and chlorine atoms together to make table salt and the relatively weak bonds between the strands of DNA that record your body's genetic code. When you stand, the normal force exerted on you by the floor arises from electric forces between charged particles in the atoms of your shoes and the atoms of the floor. The tension force in a stretched string and the adhesive force of glue are likewise due to electric interactions of atoms.

TEST YOUR UNDERSTANDING OF SECTION 21.1 Two charged objects repel each other through the electric force. The charges on the objects are (i) one positive and one negative; (ii) both positive; (iii) both negative; (iv) either (ii) or (iii); (v) any of (i), (ii), or (iii).

ANSWER

(iv) Two charged objects repel if their charges are of the same sign (either both positive or both negative).

21.2 CONDUCTORS, INSULATORS, AND INDUCED CHARGES

Some materials permit electric charge to move easily from one region of the material to another, while others do not. For example, **Fig. 21.6a** shows a copper wire supported by a nylon thread. Suppose you touch one end of the wire to a charged plastic rod and attach the other end to a metal ball that is initially uncharged; you then remove the charged rod and the wire. When you bring another charged object up close to the ball (Figs. 21.6b and 21.6c), the ball is attracted or repelled, showing that the ball has become electrically charged. Electric charge has been transferred through the copper wire between the ball and the surface of the plastic rod.

The copper wire is called a **conductor** of electricity. If you repeat the experiment using a rubber band or nylon thread in place of the wire, you find that *no* charge is transferred to the ball. These materials are called **insulators.** Conductors permit the easy movement of charge through them, while insulators do not. (The supporting nylon threads shown in Fig. 21.6 are insulators, which prevents charge from leaving the metal ball and copper wire.)

As an example, carpet fibers on a dry day are good insulators. As you walk across a carpet, the rubbing of your shoes against the fibers causes charge to build up on you, and this charge remains on you because it can't flow through the insulating fibers. If you then touch a conducting object such as a doorknob, a rapid charge transfer takes place between your finger and the doorknob, and you feel a shock. One way to prevent this is to wind some of the carpet fibers around conducting cores so that any charge that builds up on you can be transferred harmlessly to the carpet. Another solution is to coat the carpet fibers with an antistatic layer that does not easily transfer electrons to or from your shoes; this prevents any charge from building up on you in the first place.

Most metals are good conductors, while most nonmetals are insulators. Within a solid metal such as copper, one or more outer electrons in each atom become detached and can move freely throughout the material, just as the molecules of a gas can move through the spaces between the grains in a bucket of sand. The other electrons remain bound to the positively charged nuclei, which themselves are bound in nearly fixed positions within the material. In an insulator there are no, or very few, free electrons, and electric charge cannot move freely through the material. Some materials called *semiconductors* are intermediate in their properties between good conductors and good insulators.

Charging by Induction

We can charge a metal ball by using a copper wire and an electrically charged plastic rod, as in Fig. 21.6a. In this process, some of the excess electrons on the rod are transferred from it to the ball, leaving the rod with a smaller negative charge. But there is a different technique in which the plastic rod can give another object a charge of *opposite* sign without losing any of its own charge. This process is called charging by **induction.**

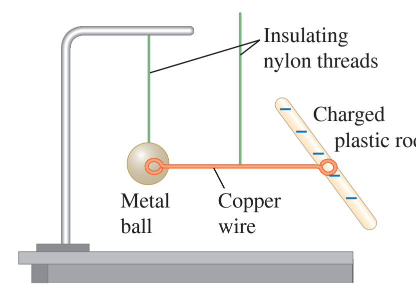
Figure 21.7 shows an example of charging by induction. An uncharged metal ball is supported on an insulating stand (Fig. 21.7a). When you bring a negatively charged rod near it, without actually touching it (Fig. 21.7b), the free electrons in the metal ball are repelled by the excess electrons on the rod, and they shift toward the right, away from the rod. They cannot escape from the ball because the supporting stand and the surrounding air are insulators. So we get excess negative charge at the right surface of the ball and a deficiency of negative charge (that is, a net positive charge) at the left surface. These excess charges are called **induced charges.**

Not all of the free electrons move to the right surface of the ball. As soon as any induced charge develops, it exerts forces toward the *left* on the other free electrons. These electrons are repelled by the negative induced charge on the right and attracted toward the positive induced charge on the left. The system reaches an equilibrium state in which the force toward the right on an electron, due to the charged rod, is just balanced by the force toward the left due to the induced charge. If we remove the charged rod, the free electrons shift back to the left, and the original neutral condition is restored.

What happens if, while the plastic rod is nearby, you touch one end of a conducting wire to the right surface of the ball and the other end to the earth (Fig. 21.7c)? The earth is a conductor, and it is so large that it can act as a practically infinite source of extra electrons or sink of unwanted electrons. Some of the negative charge flows through the wire to the earth. Now suppose you disconnect the wire (Fig. 21.7d) and then remove the rod (Fig. 21.7e); a net positive charge is left on the ball. The charge on the negatively charged rod has not changed during this process. The earth acquires a negative charge that is equal in magnitude to the induced positive charge remaining on the ball.

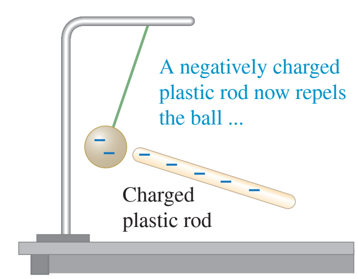
Figure **21.6** Copper is a good conductor of electricity; nylon is a good insulator. (a) The copper wire conducts charge between the metal ball and the charged plastic rod to charge the ball negatively. Afterward, the metal ball is (b) repelled by a negatively charged plastic rod and (c) attracted to a positively charged glass rod.

(a)



The wire conducts charge from the negatively charged plastic rod to the metal ball.

(b)



(c)

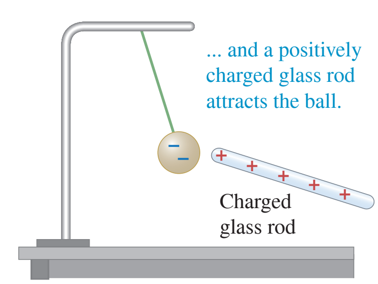
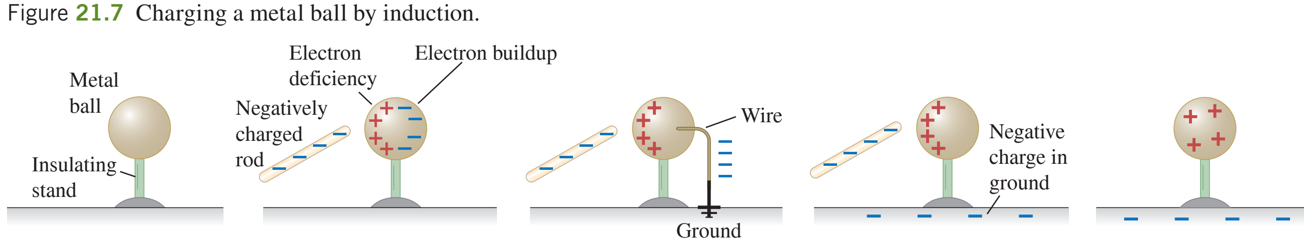


Figure **21.7** Charging a metal ball by induction.



(a) Uncharged metal ball

(b) Negative charge on rod repels electrons, creating zones of negative and positive **induced charge**.

(c) Wire lets electron buildup (induced negative charge) flow into ground.

(d) Wire removed; ball now has only an electron-deficient region of positive charge.

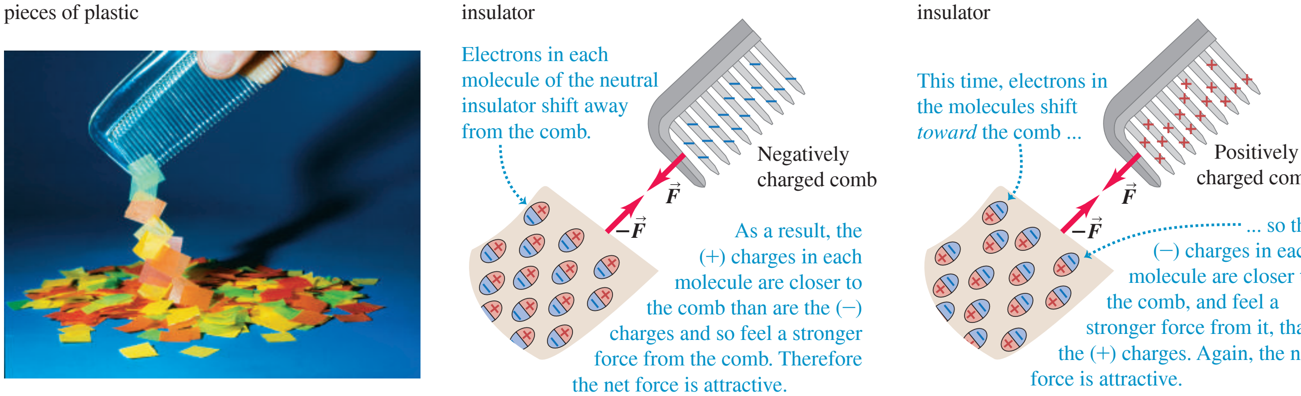
(e) Rod removed; electrons rearrange themselves, ball has overall electron deficiency (net positive charge).

Figure **21.8** The charges within the molecules of an insulating material can shift slightly. As a result, a comb with either sign of charge attracts a neutral insulator. By Newton's third law the neutral insulator exerts an equal-magnitude attractive force on the comb.

(a) A charged comb picking up uncharged pieces of plastic

(b) How a negatively charged comb attracts an insulator

(c) How a positively charged comb attracts an insulator

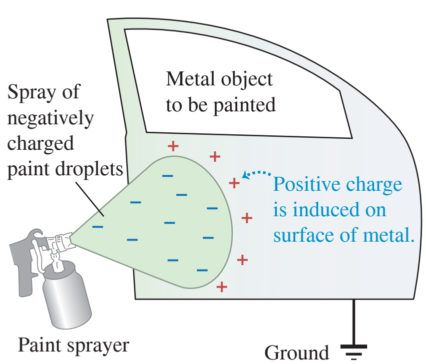


Electric Forces on Uncharged Objects

Finally, we note that a charged object can exert forces even on objects that are *not* charged themselves. If you rub a balloon on the rug and then hold the balloon against the ceiling, it sticks, even though the ceiling has no net electric charge. After you electrify a comb by running it through your hair, you can pick up uncharged bits of paper or plastic with it (**Fig. 21.8a**). How is this possible?

This interaction is an induced-charge effect. Even in an insulator, electric charges can shift back and forth a little when there is charge nearby. This is shown in Fig. 21.8b; the negatively charged plastic comb causes a slight shifting of charge within the molecules of the neutral insulator, an effect called *polarization*. The positive and negative charges in the material are present in equal amounts, but the positive charges are closer to the plastic comb and so feel an attraction that is stronger than the repulsion felt by the negative charges, giving a net attractive force. (In Section 21.3 we'll study how electric forces depend on distance.) Note that a neutral insulator is also attracted to a *positively* charged comb (Fig. 21.8c). Now the charges in the insulator shift in the opposite direction; the negative charges in the insulator are closer to the comb and feel an attractive force that is stronger than the repulsion felt by the positive charges in the insulator. Hence a charged object of *either* sign exerts an attractive force on an uncharged insulator. **Figure 21.9** shows an industrial application of this effect.

Figure **21.9** The electrostatic painting process (compare Figs. 21.7b and 21.7c). A metal object to be painted is connected to the earth ("ground"), and the paint droplets are given an electric charge as they exit the sprayer nozzle. Induced charges of the opposite sign appear in the object as the droplets approach, just as in Fig. 21.7b, and they attract the droplets to the surface. This process minimizes overspray from clouds of stray paint particles and gives a particularly smooth finish.



TEST YOUR UNDERSTANDING OF SECTION 21.2 You have two lightweight metal spheres, each hanging from an insulating nylon thread. One of the spheres has a net negative charge, while the other sphere has no net charge. (a) If the spheres are close together but do not touch, will they (i) attract each other, (ii) repel each other, or (iii) exert no force on each other? (b) You now allow the two spheres to touch. Once they have touched, will the two spheres (i) attract each other, (ii) repel each other, or (iii) exert no force on each other?

ANSWER

(a) **(i)**, (b) **(ii)** Before the two spheres touch, the negatively charged sphere exerts a repulsive force on the electrons in the other sphere, causing zones of positive and negative induced charge (see Fig. 21.7b). The positive zone is closer to the negatively charged sphere than the negative zone, so there is a net force of attraction that pulls the spheres together, like the comb and insulator in Fig. 21.8b. Once the two metal spheres touch, some of the excess electrons on the negatively charged sphere will flow onto the other sphere (because metals are conductors). Then both spheres will have a net negative charge and will repel each other.

21.3 COULOMB'S LAW

Charles Augustin de Coulomb (1736–1806) studied the interaction forces of charged particles in detail in 1784. He used a torsion balance (**Fig. 21.10a**) similar to the one used 13 years later by Cavendish to study the much weaker gravitational interaction, as we discussed in Section 13.1. For **point charges,** charged objects that are very small in comparison with the distance r between them, Coulomb found that the electric force is proportional to $1/r^2$. That is, when the distance r doubles, the force decreases to one-quarter of its initial value; when the distance is halved, the force increases to four times its initial value.

The electric force between two point charges also depends on the quantity of charge on each object, which we'll denote by q or Q. To explore this dependence, Coulomb divided a charge into two equal parts by placing a small charged spherical conductor into contact with an identical but uncharged sphere; by symmetry, the charge is shared equally between the two spheres. (Note the essential role of the principle of conservation of charge in this procedure.) Thus he could obtain one-half, one-quarter, and so on, of any initial charge. He found that the forces that two point charges q_1 and q_2 exert on each other are proportional to each charge and therefore are proportional to the *product* $q_1 q_2$ of the two charges.

Thus Coulomb established the following law:

COULOMB'S LAW The magnitude of the electric force between two point charges is directly proportional to the product of the charges and inversely proportional to the square of the distance between them.

In mathematical terms, the magnitude F of the force that each of two point charges q_1 and q_2 a distance r apart exerts on the other can be expressed as

$$F = k\frac{|q_1 q_2|}{r^2} \tag{21.1}$$

where k is a proportionality constant whose numerical value depends on the system of units used. The absolute value bars are used in Eq. (21.1) because the charges q_1 and q_2 can be either positive or negative, while the force magnitude F is always positive.

The directions of the forces the two charges exert on each other are always along the line joining them. When the charges q_1 and q_2 have the same sign, either both positive or both negative, the forces are repulsive; when the charges have opposite signs, the forces are attractive (Fig. 21.10b). The two forces obey Newton's third law; they are always equal in magnitude and opposite in direction, even when the charges are not equal in magnitude.

BIO APPLICATION Electric Forces, Sweat, and Cystic Fibrosis One way to test for the genetic disease cystic fibrosis (CF) is to measure the salt content of a person's sweat. Sweat is a mixture of water and ions, including the sodium (Na^+) and chloride (Cl^-) ions that make up ordinary salt (NaCl). When sweat is secreted by epithelial cells, some of the Cl^- ions flow from the sweat back into these cells (a process called reabsorption). The electric attraction between negative and positive charges pulls Na^+ ions along with the Cl^-. Water molecules cannot flow back into the epithelial cells, so sweat on the skin has a low salt content. However, in persons with CF the reabsorption of Cl^- ions is blocked. Hence the sweat of persons with CF is unusually salty, with up to four times the normal concentration of Cl^- and Na^+.

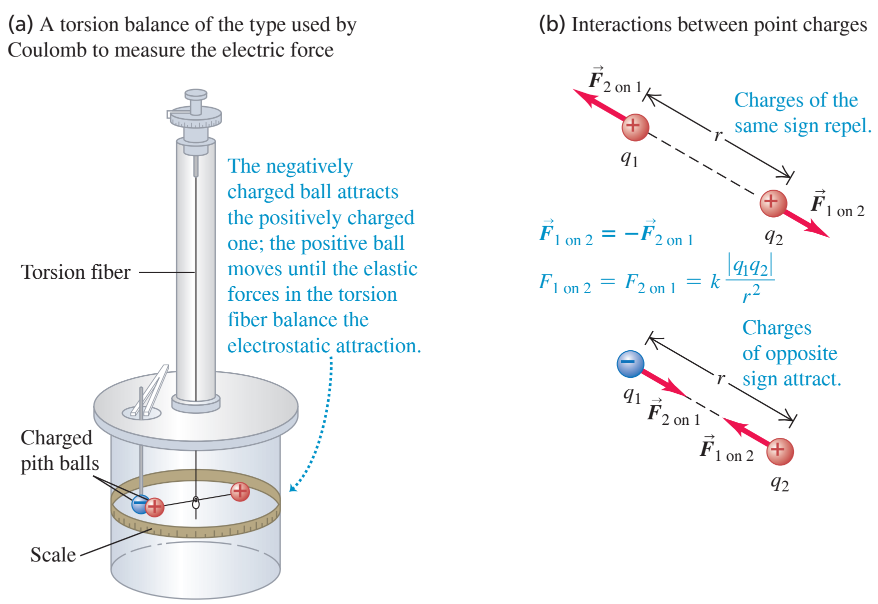


Figure **21.10** (a) Measuring the electric force between point charges. (b) The electric forces between point charges obey Newton's third law: $\vec{F}_{1\text{ on }2} = -\vec{F}_{2\text{ on }1}$.

The proportionality of the electric force to $1/r^2$ has been verified with great precision. There is no reason to suspect that the exponent is different from precisely 2. Thus the form of Eq. (21.1) is the same as that of the law of gravitation. But electric and gravitational interactions are two distinct classes of phenomena. Electric interactions depend on electric charges and can be either attractive or repulsive, while gravitational interactions depend on mass and are always attractive (because there is no such thing as negative mass).

Fundamental Electric Constants

The value of the proportionality constant k in Coulomb's law depends on the system of units used. In our study of electricity and magnetism we'll use SI units exclusively. The SI electric units include most of the familiar units such as the volt, the ampere, the ohm, and the watt. (There is *no* British system of electric units.) The SI unit of electric charge is called one **coulomb** (1 C). As of this writing (2018) the value of the constant k in Eq. (21.1) is, to nine significant figures,

$$k = 8.987551787 \times 10^9\ \mathrm{N \cdot m^2/C^2} \cong 8.988 \times 10^9\ \mathrm{N \cdot m^2/C^2}$$

Prior to 2018 k was *defined* to have a precise value in terms of the speed of light in vacuum c, whose value is in turn related to the definitions of the second and the meter that we stated in Section 1.3. (We'll see the connection between k and c in Chapter 32 when we study electromagnetic radiation.) With the revisions to the definitions of units in the International System (SI) in 2018, the value of k is no longer exact, but has a very small fractional uncertainty of about 2×10^{-10}. Because this uncertainty is so small, these revisions to the definitions of the units have no effect on the calculations that we present in this book, and don't affect the calculations that you'll do.

In principle we can measure the electric force F between two equal charges q at a measured distance r and use Coulomb's law to determine the charge. Thus we could regard the value of k as an operational definition of the coulomb. For reasons of experimental precision it is better to define the coulomb instead in terms of the most fundamental unit of charge. This is the magnitude of the charge of an electron or a proton, which is denoted by e. The value of e is *defined* to be a certain fraction of a coulomb:

$$e = 1.602176634 \times 10^{-19}\ \mathrm{C}$$

So **one coulomb (C) is that amount of charge such that $1.602176634 \times 10^{-19}$ C is equal to the magnitude of the charge on an electron or proton.** (An older definition of the coulomb involving the unit of electric *current*, or charge per unit time, was abandoned in 2018.)

One coulomb represents the negative of the total charge of about 6×10^8 electrons. For comparison, a copper cube 1 cm on a side contains about 2.4×10^{24} electrons. About 10^{19} electrons pass through the glowing filament of a flashlight bulb every second.

In SI units we usually write the constant k in Eq. (21.1) as $1/4\pi\epsilon_0$, where ϵ_0 ("epsilon-nought" or "epsilon-zero") is called the **electric constant.** This shorthand simplifies many formulas that we'll encounter in later chapters. From now on, we'll usually write Coulomb's law as

Coulomb's law:
Magnitude of electric force between two point charges

$$F = \frac{1}{4\pi\epsilon_0}\frac{|q_1 q_2|}{r^2} \qquad (21.2)$$

Values of the two charges
Distance between the two charges
Electric constant

The constants in Eq. (21.2) are approximately

$$\epsilon_0 = 8.854 \times 10^{-12}\ \mathrm{C^2/N \cdot m^2} \quad \text{and} \quad \frac{1}{4\pi\epsilon_0} = k = 8.988 \times 10^9\ \mathrm{N \cdot m^2/C^2}$$

In examples and problems we'll often use the approximate value

$$\frac{1}{4\pi\epsilon_0} = 9.0 \times 10^9\ \text{N}\cdot\text{m}^2/\text{C}^2$$

In electrostatics problems (problems that involve charges at rest), it's very unusual to encounter charges as large as 1 coulomb. Two 1 C charges separated by 1 m would exert forces on each other of magnitude 9×10^9 N (about 1 million tons)! The total charge of all the electrons in a copper one-cent coin is even greater, about 1.4×10^5 C, which shows that we can't disturb electric neutrality very much without using enormous forces. More typical values of charge range from about a microcoulomb $(1\ \mu\text{C} = 10^{-6}\ \text{C})$ to about a nanocoulomb $(1\ \text{nC} = 10^{-9}\ \text{C})$.

EXAMPLE 21.1 Electric force versus gravitational force

An α particle (the nucleus of a helium atom) has mass $m = 6.64 \times 10^{-27}$ kg and charge $q = +2e = 3.2 \times 10^{-19}$ C. Compare the magnitude of the electric repulsion between two α ("alpha") particles with that of the gravitational attraction between them.

IDENTIFY and SET UP This problem involves Newton's law for the gravitational force F_g between particles (see Section 13.1) and Coulomb's law for the electric force F_e between point charges. To compare these forces, we make our target variable the ratio F_e/F_g. We use Eq. (21.2) for F_e and Eq. (13.1) for F_g.

EXECUTE **Figure 21.11** shows our sketch. From Eqs. (21.2) and (13.1),

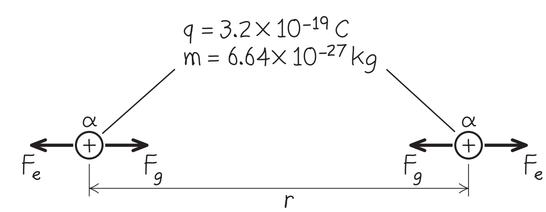
$$F_e = \frac{1}{4\pi\epsilon_0}\frac{q^2}{r^2} \qquad F_g = G\frac{m^2}{r^2}$$

These are both inverse-square forces, so the r^2 factors cancel when we take the ratio:

$$\frac{F_e}{F_g} = \frac{1}{4\pi\epsilon_0 G}\frac{q^2}{m^2}$$

$$= \frac{9.0 \times 10^9\ \text{N}\cdot\text{m}^2/\text{C}^2}{6.67 \times 10^{-11}\ \text{N}\cdot\text{m}^2/\text{kg}^2}\frac{(3.2 \times 10^{-19}\ \text{C})^2}{(6.64 \times 10^{-27}\ \text{kg})^2} = 3.1 \times 10^{35}$$

Figure 21.11 Our sketch for this problem.



EVALUATE This astonishingly large number shows that the gravitational force in this situation is completely negligible in comparison to the electric force. This is always true for interactions of atomic and subnuclear particles. But within objects the size of a person or a planet, the positive and negative charges are nearly equal in magnitude, and the net electric force is usually much *smaller* than the gravitational force.

KEYCONCEPT The magnitude of the electric force that two point charges exert on each other is (a) proportional to the product of the magnitudes of their charges and (b) inversely proportional to the square of the distance between them.

Superposition of Forces

Coulomb's law as we have stated it describes only the interaction of two *point* charges. Experiments show that when two charges exert forces simultaneously on a third charge, the total force acting on that charge is the *vector sum* of the forces that the two charges would exert individually. This important property, called the **principle of superposition of forces,** holds for any number of charges. By using this principle, we can apply Coulomb's law to *any* collection of charges. Two of the examples at the end of this section use the superposition principle.

Strictly speaking, Coulomb's law as we have stated it should be used only for point charges *in a vacuum*. If matter is present in the space between the charges, the net force acting on each charge is altered because charges are induced in the molecules of the intervening material. We'll describe this effect later. As a practical matter, though, we can use Coulomb's law unaltered for point charges in air. At normal atmospheric pressure, the presence of air changes the electric force from its vacuum value by only about one part in 2000.

PROBLEM-SOLVING STRATEGY 21.1 Coulomb's Law

IDENTIFY *the relevant concepts:* Coulomb's law describes the electric force between charged particles.

SET UP *the problem* using the following steps:

1. Sketch the locations of the charged particles and label each particle with its charge.
2. If the charges do not all lie on a single line, set up an *xy*-coordinate system.
3. The problem will ask you to find the electric force on one or more particles. Identify which these are.

EXECUTE *the solution* as follows:

1. For each particle that exerts an electric force on a given particle of interest, use Eq. (21.2) to calculate the magnitude of that force.
2. Using those magnitudes, sketch a free-body diagram showing the electric-force vectors acting on each particle of interest. The force exerted by particle 1 on particle 2 points from particle 2 toward particle 1 if the charges have opposite signs, but points from particle 2 directly away from particle 1 if the charges have the same sign.
3. Use the principle of superposition to calculate the total electric force—a *vector* sum—on each particle of interest. (Review the vector algebra in Sections 1.7 through 1.9. The method of components is often helpful.)
4. Use consistent units; SI units are completely consistent. With $1/4\pi\epsilon_0 = 9.0 \times 10^9\ \mathrm{N \cdot m^2/C^2}$, distances must be in meters, charges in coulombs, and forces in newtons.
5. Some examples and problems in this and later chapters involve *continuous* distributions of charge along a line, over a surface, or throughout a volume. In these cases the vector sum in step 3 becomes a vector *integral*. We divide the charge distribution into infinitesimal pieces, use Coulomb's law for each piece, and integrate to find the vector sum. Sometimes this can be done without actual integration.
6. Exploit any symmetries in the charge distribution to simplify your problem solving. For example, two identical charges q exert zero net electric force on a charge Q midway between them, because the forces on Q have equal magnitude and opposite direction.

EVALUATE *your answer:* Check whether your numerical results are reasonable. Confirm that the direction of the net electric force agrees with the principle that charges of the same sign repel and charges of opposite sign attract.

EXAMPLE 21.2 Force between two point charges

Two point charges, $q_1 = +25$ nC and $q_2 = -75$ nC, are separated by a distance $r = 3.0$ cm (**Fig. 21.12a**). Find the magnitude and direction of the electric force (a) that q_1 exerts on q_2 and (b) that q_2 exerts on q_1.

IDENTIFY and SET UP This problem asks for the electric forces that two charges exert on each other. We use Coulomb's law, Eq. (21.2), to calculate the magnitudes of the forces. The signs of the charges will determine the directions of the forces.

EXECUTE (a) After converting the units of r to meters and the units of q_1 and q_2 to coulombs, Eq. (21.2) gives us

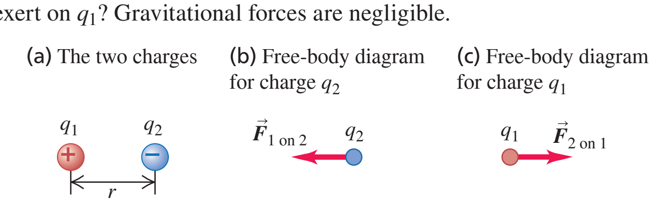
$$\begin{aligned} F_{1\text{ on }2} &= \frac{1}{4\pi\epsilon_0}\frac{|q_1 q_2|}{r^2} \\ &= (9.0 \times 10^9\ \mathrm{N \cdot m^2/C^2})\frac{|(+25 \times 10^{-9}\ \mathrm{C})(-75 \times 10^{-9}\ \mathrm{C})|}{(0.030\ \mathrm{m})^2} \\ &= 0.019\ \mathrm{N} \end{aligned}$$

The charges have opposite signs, so the force is attractive (to the left in Fig. 21.12b); that is, the force that acts on q_2 is directed toward q_1 along the line joining the two charges.

(b) Proceeding as in part (a), we have

$$F_{2\text{ on }1} = \frac{1}{4\pi\epsilon_0}\frac{|q_2 q_1|}{r^2} = F_{1\text{ on }2} = 0.019\ \mathrm{N}$$

Figure 21.12 What force does q_1 exert on q_2, and what force does q_2 exert on q_1? Gravitational forces are negligible.



The attractive force that acts on q_1 is to the right, toward q_2 (Fig. 21.12c).

EVALUATE Even though the charges have different magnitudes, the magnitude of the force that q_2 exerts on q_1 is the same as the magnitude of the force that q_1 exerts on q_2.

KEYCONCEPT The electric forces that two point charges exert on each other obey Newton's third law: The force that the first charge exerts on the second charge has the same magnitude as the force that the second charge exerts on the first charge, but points in the opposite direction. The forces are repulsive if the two charges are both positive or both negative, and attractive if one charge is positive and the other negative.

EXAMPLE 21.3 Vector addition of electric forces on a line

WITH VARIATION PROBLEMS

Two point charges are located on the x-axis of a coordinate system: $q_1 = 1.0\text{ nC}$ is at $x = +2.0\text{ cm}$, and $q_2 = -3.0\text{ nC}$ is at $x = +4.0\text{ cm}$. What is the total electric force exerted by q_1 and q_2 on a charge $q_3 = 5.0\text{ nC}$ at $x = 0$?

IDENTIFY and SET UP **Figure 21.13a** shows the situation. To find the total force on q_3, our target variable, we find the vector sum of the two electric forces on it.

EXECUTE Figure 21.13b is a free-body diagram for q_3, which is repelled by q_1 (which has the same sign) and attracted to q_2 (which has the opposite sign): $\vec{F}_{1\text{ on }3}$ is in the $-x$-direction and $\vec{F}_{2\text{ on }3}$ is in the $+x$-direction. After unit conversions, we have from Eq. (21.2)

$$F_{1\text{ on }3} = \frac{1}{4\pi\epsilon_0}\frac{|q_1q_3|}{r_{13}^2}$$

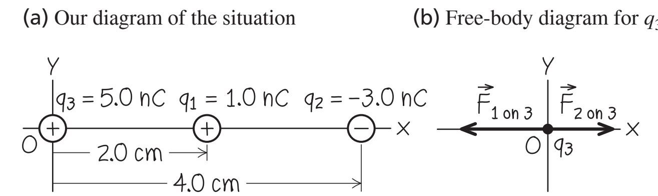
$$= (9.0 \times 10^9\text{ N}\cdot\text{m}^2/\text{C}^2)\frac{(1.0 \times 10^{-9}\text{ C})(5.0 \times 10^{-9}\text{ C})}{(0.020\text{ m})^2}$$

$$= 1.12 \times 10^{-4}\text{ N} = 112\ \mu\text{N}$$

In the same way you can show that $F_{2\text{ on }3} = 84\ \mu\text{N}$. We thus have $\vec{F}_{1\text{ on }3} = (-112\ \mu\text{N})\hat{\imath}$ and $\vec{F}_{2\text{ on }3} = (84\ \mu\text{N})\hat{\imath}$. The net force on q_3 is

$$\vec{F}_3 = \vec{F}_{1\text{ on }3} + \vec{F}_{2\text{ on }3} = (-112\ \mu\text{N})\hat{\imath} + (84\ \mu\text{N})\hat{\imath} = (-28\ \mu\text{N})\hat{\imath}$$

Figure 21.13 Our sketches for this problem.



EVALUATE As a check, note that the magnitude of q_2 is three times that of q_1, but q_2 is twice as far from q_3 as q_1. Equation (21.2) then says that $F_{2\text{ on }3}$ must be $3/2^2 = 3/4 = 0.75$ as large as $F_{1\text{ on }3}$. This agrees with our calculated values: $F_{2\text{ on }3}/F_{1\text{ on }3} = (84\ \mu\text{N})/(112\ \mu\text{N}) = 0.75$. Because $F_{2\text{ on }3}$ is the weaker force, the direction of the net force is that of $\vec{F}_{1\text{ on }3}$—that is, in the negative x-direction.

KEYCONCEPT Electric forces obey the principle of superposition of forces: The net electric force on a point charge due to two or more other charges is the vector sum of the individual electric forces that each of the other charges exert on it.

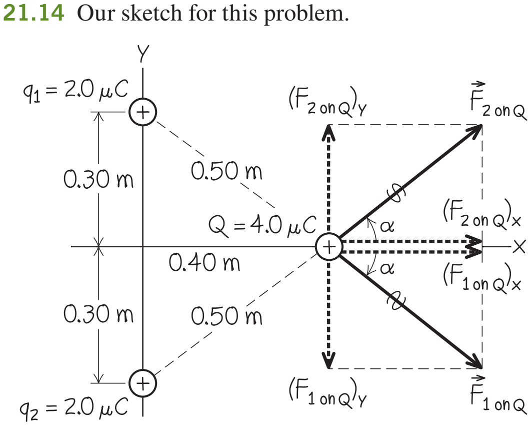
EXAMPLE 21.4 Vector addition of electric forces in a plane

WITH VARIATION PROBLEMS

Two equal positive charges $q_1 = q_2 = 2.0\ \mu\text{C}$ are located at $x = 0$, $y = 0.30\text{ m}$ and $x = 0$, $y = -0.30\text{ m}$, respectively. What are the magnitude and direction of the total electric force that q_1 and q_2 exert on a third charge $Q = 4.0\ \mu\text{C}$ at $x = 0.40\text{ m}$, $y = 0$?

IDENTIFY and SET UP As in Example 21.3, we must compute the force that each charge exerts on Q and then find the vector sum of those forces. **Figure 21.14** shows the situation. Since the three charges do not all lie on a line, the best way to calculate the forces is to use components.

Figure 21.14 Our sketch for this problem.



EXECUTE Figure 21.14 shows the forces $\vec{F}_{1\text{ on }Q}$ and $\vec{F}_{2\text{ on }Q}$ due to the identical charges q_1 and q_2, which are at equal distances from Q. From Coulomb's law, *both* forces have magnitude

$$F_{1\text{ or }2\text{ on }Q} = (9.0 \times 10^9\text{ N}\cdot\text{m}^2/\text{C}^2)\frac{(4.0 \times 10^{-6}\text{ C})(2.0 \times 10^{-6}\text{ C})}{(0.50\text{ m})^2}$$

$$= 0.29\text{ N}$$

The x-components of the two forces are equal:

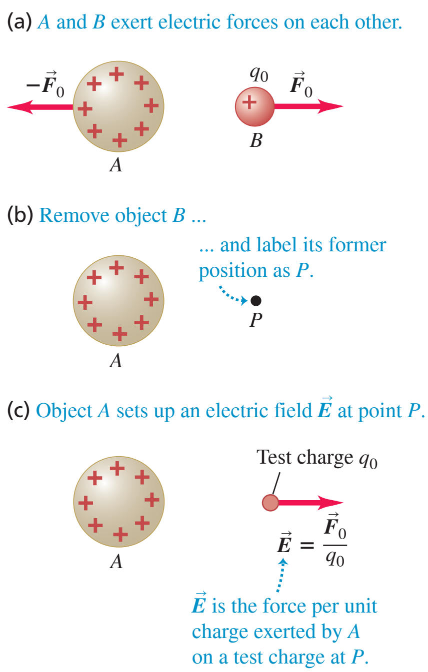
$$(F_{1\text{ or }2\text{ on }Q})_x = (F_{1\text{ or }2\text{ on }Q})\cos\alpha = (0.29\text{ N})\frac{0.40\text{ m}}{0.50\text{ m}} = 0.23\text{ N}$$

From symmetry we see that the y-components of the two forces are equal and opposite. Hence their sum is zero and the total force $\vec{F}$ on Q has only an x-component $F_x = 0.23\text{ N} + 0.23\text{ N} = 0.46\text{ N}$. The total force on Q is in the $+x$-direction, with magnitude 0.46 N.

EVALUATE The total force on Q points neither directly away from q_1 nor directly away from q_2. Rather, this direction is a compromise that points away from the *system* of charges q_1 and q_2. Can you see that the total force would *not* be in the $+x$-direction if q_1 and q_2 were not equal or if the geometrical arrangement of the changes were not so symmetric?

KEYCONCEPT When a point charge experiences electric forces due to two or more other charges, you must use vector addition to find the magnitude and direction of the net electric force on the first charge. This is most easily done by first finding the vector components of each individual electric force.

Figure **21.15** A charged object creates an electric field in the space around it.



BIO APPLICATION Sharks and the "Sixth Sense" Sharks have the ability to locate prey (such as flounder and other bottom-dwelling fish) that are completely hidden beneath the sand at the bottom of the ocean. They do this by sensing the weak electric fields produced by muscle contractions in their prey. Sharks derive their sensitivity to electric fields (a "sixth sense") from jelly-filled canals in their bodies. These canals end in pores on the shark's skin (shown in this photograph). An electric field as weak as 5×10^{-7} N/C causes charge flow within the canals and triggers a signal in the shark's nervous system. Because the shark has canals with different orientations, it can measure different components of the electric-field vector and hence determine the direction of the field.

TEST YOUR UNDERSTANDING OF SECTION 21.3 Suppose that charge q_2 in Example 21.4 were $-2.0\ \mu$C. In this case, the total electric force on Q would be (i) in the positive x-direction; (ii) in the negative x-direction; (iii) in the positive y-direction; (iv) in the negative y-direction; (v) zero; (vi) none of these.

ANSWER

(iv) The force exerted by q_1 on Q is still as in Example 21.4. The magnitude of the force exerted by q_2 on Q is still equal to F_1 on Q, but the direction of the force is now *toward* q_2 at an angle α below the x-axis. Hence the x-components of the two forces cancel while the (negative) y-components add together, and the total electric force is in the negative y-direction.

21.4 ELECTRIC FIELD AND ELECTRIC FORCES

When two electrically charged particles in empty space interact, how does each one know the other is there? We can begin to answer this question, and at the same time reformulate Coulomb's law in a very useful way, by using the concept of *electric field*.

Electric Field

To introduce this concept, let's look at the mutual repulsion of two positively charged objects A and B (**Fig. 21.15a**). Suppose B has charge q_0, and let $\vec{F}_0$ be the electric force of A on B. One way to think about this force is as an "action-at-a-distance" force—that is, as a force that acts across empty space without needing physical contact between A and B. (Gravity can also be thought of as an "action-at-a-distance" force.) But a more fruitful way to visualize the repulsion between A and B is as a two-stage process. We first envision that object A, as a result of the charge that it carries, somehow *modifies the properties of the space around it*. Then object B, as a result of the charge that *it* carries, senses how space has been modified at its position. The response of object B is to experience the force $\vec{F}_0$.

To clarify how this two-stage process occurs, we first consider object A by itself: We remove object B and label its former position as point P (Fig. 21.15b). We say that the charged object A produces or causes an **electric field** at point P (and at all other points in the neighborhood). This electric field is present at P even if there is no charge at P; it is a consequence of the charge on object A only. If a point charge q_0 is then placed at point P, it experiences the force $\vec{F}_0$. We take the point of view that this force is exerted on q_0 *by the field* at P (Fig. 21.15c). Thus the electric field is the intermediary through which A communicates its presence to q_0. Because the point charge q_0 would experience a force at *any* point in the neighborhood of A, the electric field that A produces exists at all points in the region around A.

We can likewise say that the point charge q_0 produces an electric field in the space around it and that this electric field exerts the force $-\vec{F}_0$ on object A. For each force (the force of A on q_0 and the force of q_0 on A), one charge sets up an electric field that exerts a force on the second charge. We emphasize that this is an *interaction* between *two* charged objects. A single charge produces an electric field in the surrounding space, but this electric field cannot exert a net force on the charge that created it; as we discussed in Section 4.3, an object cannot exert a net force on itself. (If this wasn't true, you would be able to lift yourself to the ceiling by pulling up on your belt!)

The electric force on a charged object is exerted by the electric field created by *other* charged objects.

To find out experimentally whether there is an electric field at a particular point, we place a small charged object, which we call a **test charge,** at the point (Fig. 21.15c). If the test charge experiences an electric force, then there is an electric field at that point. This field is produced by charges other than q_0.

Force is a vector quantity, so electric field is also a vector quantity. (Note the use of vector signs as well as boldface letters and plus, minus, and equals signs in the following discussion.) We define the *electric field* $\vec{E}$ at a point as the electric force $\vec{F}_0$

experienced by a test charge q_0 at the point, divided by the charge q_0. That is, the electric field at a certain point is equal to the *electric force per unit charge* experienced by a charge at that point:

Electric field = electric force per unit charge

$$\vec{E} = \frac{\vec{F}_0}{q_0} \qquad (21.3)$$

Electric force on a test charge q_0 due to other charges

Value of test charge

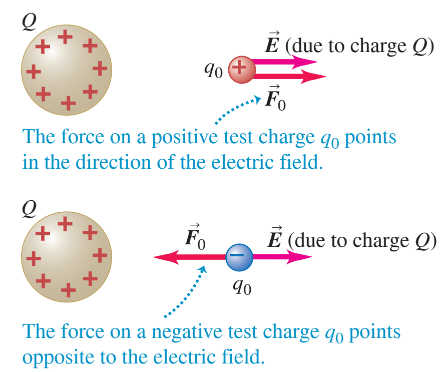
In SI units, in which the unit of force is 1 N and the unit of charge is 1 C, the unit of electric-field magnitude is 1 newton per coulomb (1 N/C).

If the field $\vec{E}$ at a certain point is known, rearranging Eq. (21.3) gives the force $\vec{F}_0$ experienced by a point charge q_0 placed at that point. This force is just equal to the electric field $\vec{E}$ produced at that point by charges other than q_0, multiplied by the charge q_0:

$$\vec{F}_0 = q_0\vec{E} \quad \text{(force exerted on a point charge } q_0 \text{ by an electric field } \vec{E}\text{)} \qquad (21.4)$$

The charge q_0 can be either positive or negative. If q_0 is *positive*, the force $\vec{F}_0$ experienced by the charge is in the same direction as $\vec{E}$; if q_0 is *negative*, $\vec{F}_0$ and $\vec{E}$ are in opposite directions (**Fig. 21.16**).

Figure **21.16** The force $\vec{F}_0 = q_0\vec{E}$ exerted on a point charge q_0 placed in an electric field $\vec{E}$.



While the electric field concept may be new to you, the basic idea—that one object sets up a field in the space around it and a second object responds to that field—is one that you've actually used before. Compare Eq. (21.4) to the familiar expression for the gravitational force $\vec{F}_g$ that the earth exerts on a mass m_0:

$$\vec{F}_g = m_0\vec{g} \qquad (21.5)$$

In this expression, $\vec{g}$ is the acceleration due to gravity. If we divide both sides of Eq. (21.5) by the mass m_0, we obtain

$$\vec{g} = \frac{\vec{F}_g}{m_0}$$

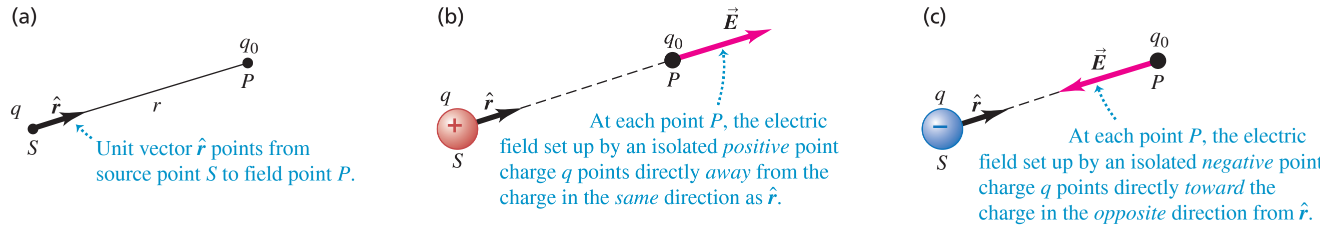
Thus $\vec{g}$ can be regarded as the gravitational force per unit mass. By analogy to Eq. (21.3), we can interpret $\vec{g}$ as the *gravitational field*. Thus we treat the gravitational interaction between the earth and the mass m_0 as a two-stage process: The earth sets up a gravitational field $\vec{g}$ in the space around it, and this gravitational field exerts a force given by Eq. (21.5) on the mass m_0 (which we can regard as a *test mass*). The gravitational field $\vec{g}$, or gravitational force per unit mass, is a useful concept because it does not depend on the mass of the object on which the gravitational force is exerted; likewise, the electric field $\vec{E}$, or electric force per unit charge, is useful because it does not depend on the charge of the object on which the electric force is exerted.

CAUTION **$\vec{F}_0 = q_0\vec{E}$ is for *point* test charges only** The electric force experienced by a test charge q_0 can vary from point to point, so the electric field can also be different at different points. For this reason, use Eq. (21.4) to find the electric force on a *point* charge only. If a charged object is large enough in size, the electric field $\vec{E}$ may be noticeably different in magnitude and direction at different points on the object, and calculating the net electric force on it can be complicated.

Electric Field of a Point Charge

If the source distribution is a point charge q, it is easy to find the electric field that it produces. We call the location of the charge the **source point,** and we call the point P where we are determining the field the **field point.** It is also useful to introduce a *unit vector* $\hat{r}$ that points along the line from source point to field point (**Fig. 21.17a**, next page). This unit vector is equal to the displacement vector $\vec{r}$ from the source point to the field point, divided by the distance $r = |\vec{r}|$ between these two

Figure **21.17** The electric field $\vec{E}$ produced at point P by an isolated point charge q at S. Note that in both (b) and (c), $\vec{E}$ is *produced* by q [see Eq. (21.7)] but *acts* on the charge q_0 at point P [see Eq. (21.4)].



points; that is, $\hat{r} = \vec{r}/r$. If we place a small test charge q_0 at the field point P, at a distance r from the source point, the magnitude F_0 of the force is given by Coulomb's law, Eq. (21.2):

$$F_0 = \frac{1}{4\pi\epsilon_0}\frac{|qq_0|}{r^2}$$

From Eq. (21.3) the magnitude E of the electric field at P is

$$E = \frac{1}{4\pi\epsilon_0}\frac{|q|}{r^2} \qquad \text{(magnitude of electric field due to a point charge)} \qquad (21.6)$$

Using the unit vector $\hat{r}$, we can write a *vector* equation that gives both the magnitude and direction of the electric field $\vec{E}$:

Electric field due to a point charge

$$\vec{E} = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}\hat{r} \qquad (21.7)$$

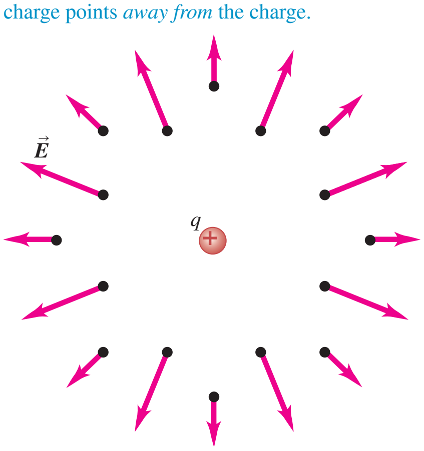
Value of point charge
Unit vector from point charge toward where field is measured
Distance from point charge to where field is measured
Electric constant

By definition, the electric field of a point charge always points *away from* a positive charge (that is, in the same direction as $\hat{r}$; see Fig. 21.17b) but *toward* a negative charge (that is, in the direction opposite $\hat{r}$; see Fig. 21.17c).

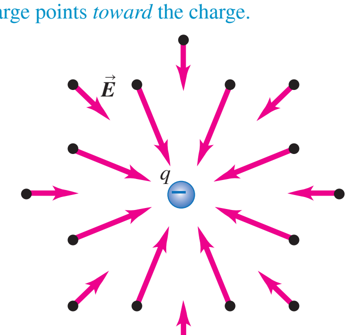
CAUTION The electric field at a point does not depend on the charge *at that* point Equation (21.7) tells you the value of the electric field $\vec{E}$ at a certain point due to a point charge q at a *different* point. The electric field at a certain point is present whether or not there is a second point charge at that point to experience the field.

Figure **21.18** A point charge q produces an electric field $\vec{E}$ at *all* points in space. The field strength decreases with increasing distance.

(a) The field produced by a positive point charge points *away from* the charge.



(b) The field produced by a negative point charge points *toward* the charge.



We have emphasized calculating the electric field $\vec{E}$ at a certain point. But since $\vec{E}$ can vary from point to point, it is not a single vector quantity but rather an *infinite* set of vector quantities, one associated with each point in space. This is an example of a **vector field. Figure 21.18** shows a number of the field vectors produced by a positive or negative point charge. If we use a rectangular (x, y, z) coordinate system, each component of $\vec{E}$ at any point is in general a function of the coordinates (x, y, z) of the point. We can represent the functions as $E_x(x, y, z)$, $E_y(x, y, z)$, and $E_z(x, y, z)$. Another example of a vector field is the velocity $\vec{v}$ of wind currents; the magnitude and direction of $\vec{v}$, and hence its vector components, vary from point to point in the atmosphere.

In some situations the magnitude and direction of the field (and hence its vector components) have the same values everywhere throughout a certain region; we then say that the field is *uniform* in this region. An important example of this is the electric field inside a *conductor*. If there is an electric field within a conductor, the field exerts a force on every charge in the conductor, giving the free charges a net motion. By definition an electrostatic situation is one in which the charges have *no* net motion. We conclude that *in electrostatics the electric field at every point within the material of a conductor must be zero.* (Note that we are not saying that the field is necessarily zero in a *hole* inside a conductor.)

In summary, our description of electric interactions has two parts. First, a given charge distribution acts as a source of electric field. Second, the electric field exerts a force on any charge that is present in the field. Our analysis often has two corresponding steps: first, calculating the field caused by a source charge distribution; second, looking at the effect of the field in terms of force and motion. The second step often involves Newton's laws as well as the principles of electric interactions. In the next section we show how to calculate fields caused by various source distributions, but first here are three examples of calculating the field due to a point charge and of finding the force on a charge due to a given field $\vec{E}$.

EXAMPLE 21.5 Electric-field magnitude for a point charge

What is the magnitude of the electric field $\vec{E}$ at a field point 2.0 m from a point charge $q = 4.0$ nC?

IDENTIFY and SET UP This problem concerns the electric field due to a point charge. We are given the magnitude of the charge and the distance from the charge to the field point, so we use Eq. (21.6) to calculate the field magnitude E.

EXECUTE From Eq. (21.6),

$$E = \frac{1}{4\pi\epsilon_0}\frac{|q|}{r^2} = (9.0 \times 10^9\ \mathrm{N\cdot m^2/C^2})\frac{4.0 \times 10^{-9}\ \mathrm{C}}{(2.0\ \mathrm{m})^2} = 9.0\ \mathrm{N/C}$$

EVALUATE Our result $E = 9.0$ N/C means that if we placed a 1.0 C charge at a point 2.0 m from q, it would experience a 9.0 N force. The force on a 2.0 C charge at that point would be $(2.0\ \mathrm{C})(9.0\ \mathrm{N/C}) = 18$ N, and so on.

KEYCONCEPT The magnitude of the electric field that a point charge produces at a given point is (a) proportional to the magnitude of the charge and (b) inversely proportional to the square of the distance from the point charge to the point where the field is measured.

EXAMPLE 21.6 Electric-field vector for a point charge

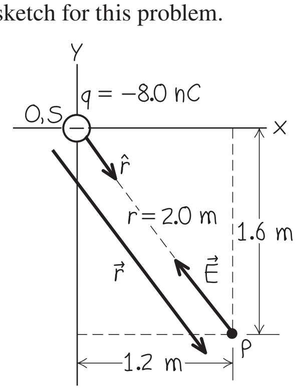
A point charge $q = -8.0$ nC is located at the origin. Find the electric-field vector at the field point $x = 1.2$ m, $y = -1.6$ m.

IDENTIFY and SET UP We must find the electric-field *vector* $\vec{E}$ due to a point charge. **Figure 21.19** shows the situation. We use Eq. (21.7); to do this, we must find the distance r from the source point S (the position of the charge q, which in this example is at the origin O) to the field point P, and we must obtain an expression for the unit vector $\hat{r} = \vec{r}/r$ that points from S to P.

EXECUTE The distance from S to P is

$$r = \sqrt{x^2 + y^2} = \sqrt{(1.2\ \mathrm{m})^2 + (-1.6\ \mathrm{m})^2} = 2.0\ \mathrm{m}$$

Figure **21.19** Our sketch for this problem.



The unit vector $\hat{r}$ is then

$$\hat{r} = \frac{\vec{r}}{r} = \frac{x\hat{\imath} + y\hat{\jmath}}{r}$$

$$= \frac{(1.2\ \mathrm{m})\hat{\imath} + (-1.6\ \mathrm{m})\hat{\jmath}}{2.0\ \mathrm{m}} = 0.60\hat{\imath} - 0.80\hat{\jmath}$$

Then, from Eq. (21.7),

$$\vec{E} = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}\hat{r}$$

$$= (9.0 \times 10^9\ \mathrm{N\cdot m^2/C^2})\frac{(-8.0 \times 10^{-9}\ \mathrm{C})}{(2.0\ \mathrm{m})^2}(0.60\hat{\imath} - 0.80\hat{\jmath})$$

$$= (-11\ \mathrm{N/C})\hat{\imath} + (14\ \mathrm{N/C})\hat{\jmath}$$

EVALUATE Since q is negative, $\vec{E}$ points from the field point to the charge (the source point), in the direction opposite to $\hat{r}$ (compare Fig. 21.17c). We leave the calculation of the magnitude and direction of $\vec{E}$ to you.

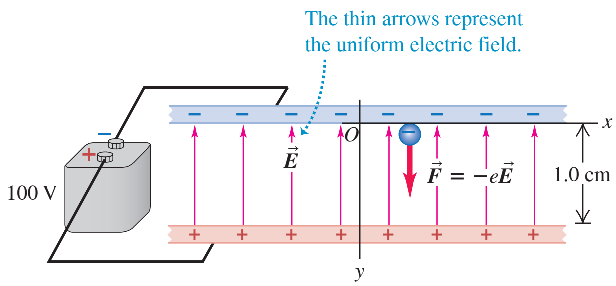
KEYCONCEPT When you calculate the electric field due to a point charge q, the unit vector $\hat{r}$ always points from the charge to the point where the field is measured. If q is positive, the field is in the same direction as $\hat{r}$ (away from the charge); if q is negative, the field is in the direction opposite to $\hat{r}$ (toward the charge.)

EXAMPLE 21.7 Electron in a uniform field

When the terminals of a battery are connected to two parallel conducting plates with a small gap between them, the resulting charges on the plates produce a nearly uniform electric field $\vec{E}$ between the plates. (In the next section we'll see why this is.) If the plates are 1.0 cm apart and are connected to a 100 volt battery as shown in **Fig. 21.20** (next page), the field is vertically upward and has magnitude $E = 1.00 \times 10^4$ N/C. (a) If an electron (charge $-e = -1.60 \times 10^{-9}$ C, mass $m = 9.11 \times 10^{-31}$ kg) is released from rest at the upper plate, what is its acceleration? (b) What speed and kinetic energy does it acquire while traveling 1.0 cm to the lower plate? (c) How long does it take to travel this distance?

Continued

Figure **21.20** A uniform electric field between two parallel conducting plates connected to a 100 volt battery. (The separation of the plates is exaggerated in this figure relative to the dimensions of the plates.)



IDENTIFY and SET UP This example involves the relationship between electric field and electric force. It also involves the relationship between force and acceleration, the definition of kinetic energy, and the kinematic relationships among acceleration, distance, velocity, and time. Figure 21.20 shows our coordinate system. We are given the electric field, so we use Eq. (21.4) to find the force on the electron and Newton's second law to find its acceleration. Because the field is uniform, the force is constant and we can use the constant-acceleration formulas from Chapter 2 to find the electron's velocity and travel time. We find the kinetic energy from $K = \frac{1}{2}mv^2$.

EXECUTE (a) Although $\vec{E}$ is upward (in the $+y$-direction), $\vec{F}$ is downward (because the electron's charge is negative) and so F_y is negative. Because F_y is constant, the electron's acceleration is constant:

$$a_y = \frac{F_y}{m} = \frac{-eE}{m} = \frac{(-1.60 \times 10^{-19}\ \text{C})(1.00 \times 10^4\ \text{N/C})}{9.11 \times 10^{-31}\ \text{kg}}$$
$$= -1.76 \times 10^{15}\ \text{m/s}^2$$

(b) The electron starts from rest, so its motion is in the y-direction only (the direction of the acceleration). We can find the electron's speed at any position y from the constant-acceleration Eq. (2.13), $v_y^2 = v_{0y}^2 + 2a_y(y - y_0)$. We have $v_{0y} = 0$ and $y_0 = 0$, so at $y = -1.0$ cm $= -1.0 \times 10^{-2}$ m we have

$$|v_y| = \sqrt{2a_y y} = \sqrt{2(-1.76 \times 10^{15}\ \text{m/s}^2)(-1.0 \times 10^{-2}\ \text{m})}$$
$$= 5.9 \times 10^6\ \text{m/s}$$

The velocity is downward, so $v_y = -5.9 \times 10^6$ m/s. The electron's kinetic energy is

$$K = \tfrac{1}{2}mv^2 = \tfrac{1}{2}(9.11 \times 10^{-31}\ \text{kg})(5.9 \times 10^6\ \text{m/s})^2$$
$$= 1.6 \times 10^{-17}\ \text{J}$$

(c) From Eq. (2.8) for constant acceleration, $v_y = v_{0y} + a_y t$, so

$$t = \frac{v_y - v_{0y}}{a_y} = \frac{(-5.9 \times 10^6\ \text{m/s}) - (0)}{-1.76 \times 10^{15}\ \text{m/s}^2}$$
$$= 3.4 \times 10^{-9}\ \text{s}$$

EVALUATE Our results show that in problems concerning subatomic particles such as electrons, many quantities—including acceleration, speed, kinetic energy, and time—will have *very* different values from those typical of everyday objects such as baseballs and automobiles.

KEYCONCEPT If you place a point charge q_0 at a location where the electric field is $\vec{E}$, the electric force on the point charge is $q_0\vec{E}$. This force is in the same direction as $\vec{E}$ if q_0 is positive; the force is in the direction opposite to $\vec{E}$ if q_0 is negative.

TEST YOUR UNDERSTANDING OF SECTION 21.4 (a) A negative point charge moves along a straight-line path directly toward a stationary positive point charge. Which aspect(s) of the electric force on the negative point charge will remain constant as it moves? (i) Magnitude; (ii) direction; (iii) both magnitude and direction; (iv) neither magnitude nor direction. (b) A negative point charge moves along a circular orbit around a positive point charge. Which aspect(s) of the electric force on the negative point charge will remain constant as it moves? (i) Magnitude; (ii) direction; (iii) both magnitude and direction; (iv) neither magnitude nor direction.

ANSWER

(a) **(ii)**, (b) **(i)** The electric field $\vec{E}$ produced by a positive point charge points directly away from the charge (see Fig. 21.18a) and has a magnitude that depends on the distance r from the charge to the field point. Hence a second, negative point charge $q < 0$ will feel a force $\vec{F} = q\vec{E}$ that points directly toward the positive charge and has a magnitude that depends on the distance r between the two charges. If the negative charge moves directly toward the positive charge, the direction of the force remains the same but the force magnitude increases as the distance r decreases. If the negative charge moves in a circle around the positive charge, the force magnitude stays the same (because the distance r is constant) but the force direction changes.

21.5 ELECTRIC-FIELD CALCULATIONS

Equation (21.7) gives the electric field caused by a single point charge. But in most realistic situations that involve electric fields and forces, we encounter charge that is *distributed* over space. The charged plastic and glass rods in Fig. 21.1 have electric charge distributed over their surfaces, as does the imaging drum of a laser printer (Fig. 21.2). In this section we'll learn to calculate electric fields caused by various distributions of electric charge. Calculations of this kind are of tremendous importance for technological applications of electric forces. To determine the trajectories of atomic nuclei in an accelerator for cancer radiotherapy or of charged particles in a semiconductor electronic device, you have to know the detailed nature of the electric field acting on the charges.

The Superposition of Electric Fields

To find the field caused by a charge distribution, we imagine the distribution to be made up of many point charges $q_1, q_2, q_3, \ldots$. (This is actually quite a realistic description, since we have seen that charge is carried by electrons and protons that are so small as to be almost pointlike.) At any given point P, each point charge produces its own electric field $\vec{E}_1, \vec{E}_2, \vec{E}_3, \ldots$, so a test charge q_0 placed at P experiences a force $\vec{F}_1 = q_0\vec{E}_1$ from charge q_1, a force $\vec{F}_2 = q_0\vec{E}_2$ from charge q_2, and so on. From the principle of superposition of forces discussed in Section 21.3, the *total* force $\vec{F}_0$ that the charge distribution exerts on q_0 is the vector sum of these individual forces:

$$\vec{F}_0 = \vec{F}_1 + \vec{F}_2 + \vec{F}_3 + \cdots = q_0\vec{E}_1 + q_0\vec{E}_2 + q_0\vec{E}_3 + \cdots$$

The combined effect of all the charges in the distribution is described by the *total* electric field $\vec{E}$ at point P. From the definition of electric field, Eq. (21.3), this is

$$\vec{E} = \frac{\vec{F}_0}{q_0} = \vec{E}_1 + \vec{E}_2 + \vec{E}_3 + \cdots$$

This equation is the mathematical expression of the *principle of superposition of electric fields,* illustrated in **Fig. 21.21**:

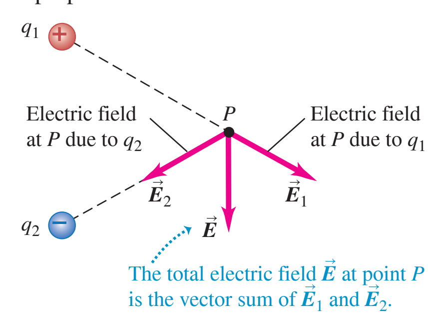


Figure **21.21** Illustrating the principle of superposition of electric fields.

PRINCIPLE OF SUPERPOSITION OF ELECTRIC FIELDS The total electric field at a point P is the vector sum of the fields at P due to each point charge in the charge distribution.

When charge is distributed along a line, over a surface, or through a volume, a few additional terms are useful. For a line charge distribution (such as a long, thin, charged plastic rod), we use λ (the Greek letter lambda) to represent the **linear charge density** (charge per unit length, measured in C/m). When charge is distributed over a surface (such as the surface of the imaging drum of a laser printer), we use σ (sigma) to represent the **surface charge density** (charge per unit area, measured in $\mathrm{C/m^2}$). And when charge is distributed through a volume, we use ρ (rho) to represent the **volume charge density** (charge per unit volume, $\mathrm{C/m^3}$).

Some of the calculations in the following examples may look complex. After you've worked through the examples one step at a time, the process will seem less formidable. We'll use many of the calculational techniques in these examples in Chapter 28 to calculate the *magnetic* fields caused by charges in motion.

PROBLEM-SOLVING STRATEGY 21.2 Electric-Field Calculations

IDENTIFY *the relevant concepts:* Use the principle of superposition to calculate the electric field due to a discrete or continuous charge distribution.

SET UP *the problem* using the following steps:

1. Make a drawing showing the locations of the charges and your choice of coordinate axes.
2. On your drawing, indicate the position of the *field point P* (the point at which you want to calculate the electric field $\vec{E}$).

EXECUTE *the solution* as follows:

1. Use consistent units. Distances must be in meters and charge must be in coulombs. If you are given centimeters or nanocoulombs, don't forget to convert.
2. Distinguish between the source point S and the field point P. The field produced by a point charge always points from S to P if the charge is positive, and from P to S if the charge is negative.
3. Use *vector* addition when applying the principle of superposition; review the treatment of vector addition in Chapter 1 if necessary.
4. Simplify your calculations by exploiting any symmetries in the charge distribution.
5. If the charge distribution is continuous, define a small element of charge that can be considered as a point, find its electric field at P, and find a way to add the fields of all the charge elements by doing an integral. Usually it is easiest to do this for each component of $\vec{E}$ separately, so you may need to evaluate more than one integral. Ensure that the limits on your integrals are correct; especially when the situation has symmetry, don't count a charge twice.

EVALUATE *your answer:* Check that the direction of $\vec{E}$ is reasonable. If your result for the electric-field magnitude E is a function of position (say, the coordinate x), check your result in any limits for which you know what the magnitude should be. When possible, check your answer by calculating it in a different way.

EXAMPLE 21.8 Field of an electric dipole

WITH VARIATION PROBLEMS

Point charges $q_1 = +12\text{ nC}$ and $q_2 = -12\text{ nC}$ are 0.100 m apart (**Fig. 21.22**). (Such pairs of point charges with equal magnitude and opposite sign are called *electric dipoles*.) Compute the electric field caused by q_1, the field caused by q_2, and the total field (a) at point a; (b) at point b; and (c) at point c.

IDENTIFY and SET UP We must find the total electric field at various points due to two point charges. We use the principle of superposition: $\vec{E} = \vec{E}_1 + \vec{E}_2$. Figure 21.22 shows the coordinate system and the locations of the field points a, b, and c.

EXECUTE At each field point, $\vec{E}$ depends on $\vec{E}_1$ and $\vec{E}_2$ there; we first calculate the magnitudes E_1 and E_2 at each field point. At a the magnitude of the field $\vec{E}_{1a}$ caused by q_1 is

$$\begin{aligned}E_{1a} &= \frac{1}{4\pi\epsilon_0}\frac{|q_1|}{r^2}\\ &= (9.0\times10^9\text{ N}\cdot\text{m}^2/\text{C}^2)\frac{12\times10^{-9}\text{ C}}{(0.060\text{ m})^2}\\ &= 3.0\times10^4\text{ N/C}\end{aligned}$$

We calculate the other field magnitudes in a similar way. The results are

$$\begin{aligned}E_{1a} &= 3.0\times10^4\text{ N/C}\\ E_{1b} &= 6.8\times10^4\text{ N/C}\\ E_{1c} &= 6.39\times10^3\text{ N/C}\\ E_{2a} &= 6.8\times10^4\text{ N/C}\\ E_{2b} &= 0.55\times10^4\text{ N/C}\\ E_{2c} &= E_{1c} = 6.39\times10^3\text{ N/C}\end{aligned}$$

The *directions* of the corresponding fields are in all cases *away* from the positive charge q_1 and *toward* the negative charge q_2.

(a) At a, $\vec{E}_{1a}$ and $\vec{E}_{2a}$ are both directed to the right, so

$$\vec{E}_a = E_{1a}\hat{\imath} + E_{2a}\hat{\imath} = (9.8\times10^4\text{ N/C})\hat{\imath}$$

(b) At b, $\vec{E}_{1b}$ is directed to the left and $\vec{E}_{2b}$ is directed to the right, so

$$\vec{E}_b = -E_{1b}\hat{\imath} + E_{2b}\hat{\imath} = (-6.2\times10^4\text{ N/C})\hat{\imath}$$

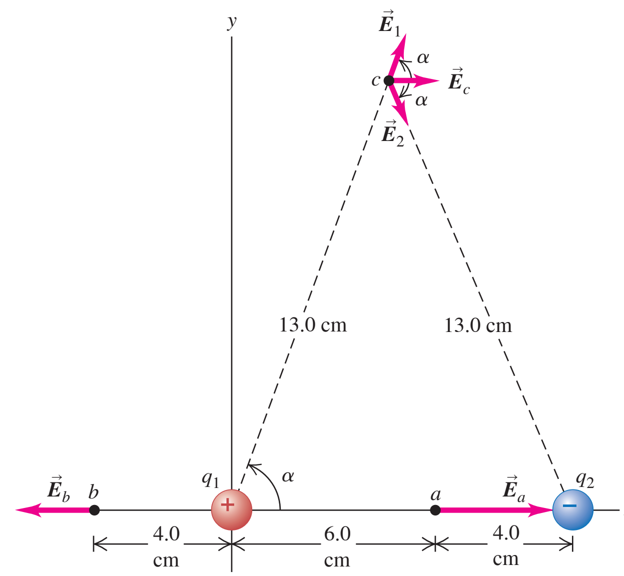
(c) Figure 21.22 shows the directions of $\vec{E}_1$ and $\vec{E}_2$ at c. Both vectors have the same x-component:

$$\begin{aligned}E_{1cx} = E_{2cx} = E_{1c}\cos\alpha &= (6.39\times10^3\text{ N/C})\left(\tfrac{5}{13}\right)\\ &= 2.46\times10^3\text{ N/C}\end{aligned}$$

From symmetry, E_{1y} and E_{2y} are equal and opposite, so their sum is zero. Hence

$$\vec{E}_c = 2(2.46\times10^3\text{ N/C})\hat{\imath} = (4.9\times10^3\text{ N/C})\hat{\imath}$$

Figure 21.22 Electric field at three points, a, b, and c, set up by charges q_1 and q_2, which form an electric dipole.



EVALUATE We can also find $\vec{E}_c$ by using Eq. (21.7) for the field of a point charge. The displacement vector $\vec{r}_1$ from q_1 to point c is $\vec{r}_1 = r\cos\alpha\hat{\imath} + r\sin\alpha\hat{\jmath}$. Hence the unit vector that points from q_1 to point c is $\hat{r}_1 = \vec{r}_1/r = \cos\alpha\hat{\imath} + \sin\alpha\hat{\jmath}$. By symmetry, the unit vector that points from q_2 to point c has the opposite x-component but the same y-component: $\hat{r}_2 = -\cos\alpha\hat{\imath} + \sin\alpha\hat{\jmath}$. We can now use Eq. (21.7) to write the fields $\vec{E}_{1c}$ and $\vec{E}_{2c}$ at c in vector form, then find their sum. Since $q_2 = -q_1$ and the distance r to c is the same for both charges,

$$\begin{aligned}\vec{E}_c = \vec{E}_{1c} + \vec{E}_{2c} &= \frac{1}{4\pi\epsilon_0}\frac{q_1}{r^2}\hat{r}_1 + \frac{1}{4\pi\epsilon_0}\frac{q_2}{r^2}\hat{r}_2\\ &= \frac{1}{4\pi\epsilon_0 r^2}(q_1\hat{r}_1 + q_2\hat{r}_2)\\ &= \frac{q_1}{4\pi\epsilon_0 r^2}(\hat{r}_1 - \hat{r}_2)\\ &= \frac{1}{4\pi\epsilon_0}\frac{q_1}{r^2}(2\cos\alpha\hat{\imath})\\ &= 2(9.0\times10^9\text{ N}\cdot\text{m}^2/\text{C}^2)\frac{12\times10^{-9}\text{ C}}{(0.13\text{ m})^2}\left(\tfrac{5}{13}\right)\hat{\imath}\\ &= (4.9\times10^3\text{ N/C})\hat{\imath}\end{aligned}$$

This is the same as we calculated in part (c).

KEYCONCEPT To find the net electric field at a point due to two or more point charges, first find the field at that point due to each individual charge. Then use vector addition to find the magnitude and direction of the net field at that point.

EXAMPLE 21.9 Field of a ring of charge

WITH VARIATION PROBLEMS

Charge Q is uniformly distributed around a conducting ring of radius a (**Fig. 21.23**). Find the electric field at a point P on the ring axis at a distance x from its center.

IDENTIFY and SET UP This is a problem in the superposition of electric fields. Each bit of charge around the ring produces an electric field at an arbitrary point on the x-axis; our target variable is the total field at this point due to all such bits of charge.

EXECUTE We divide the ring into infinitesimal segments ds as shown in Fig. 21.23. In terms of the linear charge density $\lambda = Q/2\pi a$, the charge in a segment of length ds is $dQ = \lambda\, ds$. Consider two identical segments, one as shown in the figure at $y = a$ and another halfway around the ring at $y = -a$. From Example 21.4, we see that the net force $d\vec{F}$ they exert on a point test charge at P, and thus their net field $d\vec{E}$, are directed along the x-axis. The same is true for any such pair of segments around the ring, so the *net* field at P is along the x-axis: $\vec{E} = E_x\hat{\imath}$.

To calculate E_x, note that the square of the distance r from a single ring segment to the point P is $r^2 = x^2 + a^2$. Hence the magnitude of this segment's contribution $d\vec{E}$ to the electric field at P is

$$dE = \frac{1}{4\pi\epsilon_0}\frac{dQ}{x^2 + a^2}$$

The x-component of this field is $dE_x = dE\cos\alpha$. We know $dQ = \lambda\, ds$ and Fig. 21.23 shows that $\cos\alpha = x/r = x/(x^2 + a^2)^{1/2}$, so

$$dE_x = dE\cos\alpha = \frac{1}{4\pi\epsilon_0}\frac{dQ}{x^2 + a^2}\frac{x}{\sqrt{x^2 + a^2}} = \frac{1}{4\pi\epsilon_0}\frac{\lambda x}{(x^2 + a^2)^{3/2}}ds$$

To find E_x we integrate this expression over the entire ring—that is, for s from 0 to $2\pi a$ (the circumference of the ring). The integrand has the same value for all points on the ring, so it can be taken outside the integral. Hence we get

$$E_x = \int dE_x = \frac{1}{4\pi\epsilon_0}\frac{\lambda x}{(x^2 + a^2)^{3/2}}\int_0^{2\pi a} ds$$

$$= \frac{1}{4\pi\epsilon_0}\frac{\lambda x}{(x^2 + a^2)^{3/2}}(2\pi a)$$

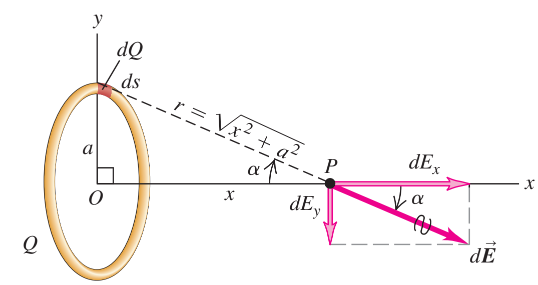


Figure 21.23 Calculating the electric field on the axis of a ring of charge. In this figure, the charge is assumed to be positive.

$$\vec{E} = E_x\hat{\imath} = \frac{1}{4\pi\epsilon_0}\frac{Qx}{(x^2 + a^2)^{3/2}}\hat{\imath} \tag{21.8}$$

EVALUATE Equation (21.8) shows that $\vec{E} = \mathbf{0}$ at the center of the ring ($x = 0$). This makes sense; charges on opposite sides of the ring push in opposite directions on a test charge at the center, and the vector sum of each such pair of forces is zero. When the field point P is much farther from the ring than the ring's radius, we have $x \gg a$ and the denominator in Eq. (21.8) becomes approximately equal to x^3. In this limit the electric field at P is

$$\vec{E} = \frac{1}{4\pi\epsilon_0}\frac{Q}{x^2}\hat{\imath}$$

That is, when the ring is so far away that its radius is negligible in comparison to the distance x, its field is the same as that of a point charge.

KEYCONCEPT To find the vector components of the net electric field at a point due to a continuous distribution of charge, first divide the distribution into infinitesimally small segments. Then find the components of the field at the point due to one such segment. Finally, integrate each component of the field due to a segment over all segments in the charge distribution.

EXAMPLE 21.10 Field of a charged line segment

WITH VARIATION PROBLEMS

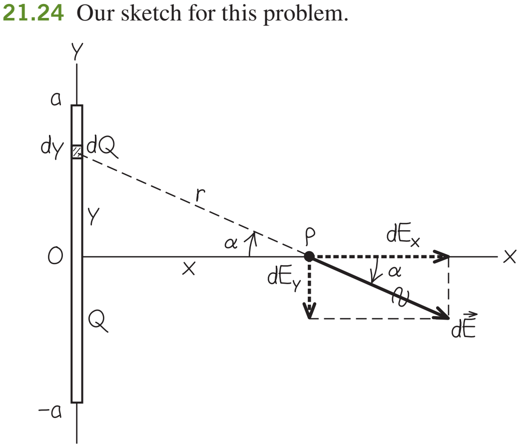
Positive charge Q is distributed uniformly along the y-axis between $y = -a$ and $y = +a$. Find the electric field at point P on the x-axis at a distance x from the origin.

IDENTIFY and SET UP **Figure 21.24** shows the situation. As in Example 21.9, we must find the electric field due to a continuous distribution of charge. Our target variable is an expression for the electric field at P as a function of x. The x-axis is a perpendicular bisector of the segment, so we can use a symmetry argument.

EXECUTE We divide the line charge of length $2a$ into infinitesimal segments of length dy. The linear charge density is $\lambda = Q/2a$, and the charge in a segment is $dQ = \lambda\, dy = (Q/2a)dy$. The distance r from a segment at height y to the field point P is $r = (x^2 + y^2)^{1/2}$, so the magnitude of the field at P due to the segment at height y is

$$dE = \frac{1}{4\pi\epsilon_0}\frac{dQ}{r^2} = \frac{1}{4\pi\epsilon_0}\frac{Q}{2a}\frac{dy}{(x^2 + y^2)}$$

Figure 21.24 Our sketch for this problem.



Continued

Figure 21.24 shows that the x- and y-components of this field are $dE_x = dE\cos\alpha$ and $dE_y = -dE\sin\alpha$, where $\cos\alpha = x/r$ and $\sin\alpha = y/r$. Hence

$$dE_x = \frac{1}{4\pi\epsilon_0}\frac{Q}{2a}\frac{x\,dy}{(x^2+y^2)^{3/2}}$$

$$dE_y = -\frac{1}{4\pi\epsilon_0}\frac{Q}{2a}\frac{y\,dy}{(x^2+y^2)^{3/2}}$$

To find the total field at P, we must sum the fields from all segments along the line—that is, we must integrate from $y = -a$ to $y = +a$. You should work out the details of the integration (a table of integrals will help). The results are

$$E_x = \frac{1}{4\pi\epsilon_0}\frac{Q}{2a}\int_{-a}^{+a}\frac{x\,dy}{(x^2+y^2)^{3/2}} = \frac{Q}{4\pi\epsilon_0}\frac{1}{x\sqrt{x^2+a^2}}$$

$$E_y = -\frac{1}{4\pi\epsilon_0}\frac{Q}{2a}\int_{-a}^{+a}\frac{y\,dy}{(x^2+y^2)^{3/2}} = 0$$

or, in vector form,

$$\vec{E} = \frac{1}{4\pi\epsilon_0}\frac{Q}{x\sqrt{x^2+a^2}}\hat{\imath} \qquad (21.9)$$

$\vec{E}$ points away from the line of charge if λ is positive and toward the line of charge if λ is negative.

EVALUATE Using a symmetry argument as in Example 21.9, we could have guessed that E_y would be zero; if we place a positive test charge at P, the upper half of the line of charge pushes downward on it, and the lower half pushes up with equal magnitude. Symmetry also tells us that the upper and lower halves of the segment contribute equally to the total field at P.

If the segment is very *short* (or the field point is very far from the segment) so that $x \gg a$, we can ignore a in the denominator of Eq. (21.9). Then the field becomes that of a point charge, just as in Example 21.9:

$$\vec{E} = \frac{1}{4\pi\epsilon_0}\frac{Q}{x^2}\hat{\imath}$$

To see what happens if the segment is very *long* (or the field point is very close to it) so that $a \gg x$, we first rewrite Eq. (21.9) slightly:

$$\vec{E} = \frac{1}{2\pi\epsilon_0}\frac{\lambda}{x\sqrt{(x^2/a^2)+1}}\hat{\imath} \qquad (21.10)$$

In the limit $a \gg x$ we can ignore x^2/a^2 in the denominator of Eq. (21.10), so

$$\vec{E} = \frac{\lambda}{2\pi\epsilon_0 x}\hat{\imath}$$

This is the field of an *infinitely long* line of charge. At any point P at a perpendicular distance r from the line in *any* direction, $\vec{E}$ has magnitude

$$E = \frac{\lambda}{2\pi\epsilon_0 r} \qquad \text{(infinite line of charge)}$$

Note that this field is proportional to $1/r$ rather than to $1/r^2$ as for a point charge.

There's really no such thing in nature as an infinite line of charge. But when the field point is close enough to the line, there's very little difference between the result for an infinite line and the real-life finite case. For example, if the distance r of the field point from the center of the line is 1% of the length of the line, the value of E differs from the infinite-length value by less than 0.02%.

KEYCONCEPT The electric field due to a symmetrical distribution of charge is most easily calculated at a point of symmetry. Whenever possible, take advantage of the symmetry of the situation to check your results.

EXAMPLE 21.11 Field of a uniformly charged disk

A nonconducting disk of radius R has a uniform positive surface charge density σ. Find the electric field at a point along the axis of the disk a distance x from its center. Assume that x is positive.

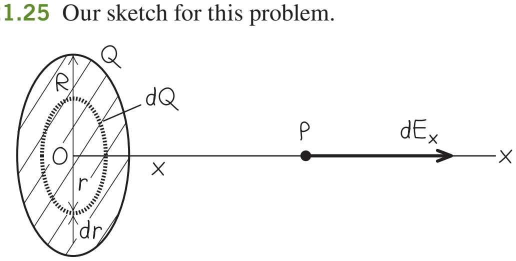
IDENTIFY and SET UP **Figure 21.25** shows the situation. We represent the charge distribution as a collection of concentric rings of charge dQ. In Example 21.9 we obtained Eq. (21.8) for the field on the axis of a single uniformly charged ring, so all we need do here is integrate the contributions of our rings.

EXECUTE A typical ring has charge dQ, inner radius r, and outer radius $r + dr$. Its area is approximately equal to its width dr times its circumference $2\pi r$, or $dA = 2\pi r\,dr$. The charge per unit area is $\sigma = dQ/dA$, so the charge of the ring is $dQ = \sigma\,dA = 2\pi\sigma r\,dr$. We use dQ in place of Q in Eq. (21.8), the expression for the field due to a ring that we found in Example 21.9, and replace the ring radius a with r. Then the field component dE_x at point P due to this ring is

$$dE_x = \frac{1}{4\pi\epsilon_0}\frac{2\pi\sigma r x\,dr}{(x^2+r^2)^{3/2}}$$

To find the total field due to all the rings, we integrate dE_x over r from $r = 0$ to $r = R$ (*not* from $-R$ to R):

Figure **21.25** Our sketch for this problem.



$$E_x = \int_0^R \frac{1}{4\pi\epsilon_0}\frac{(2\pi\sigma r\,dr)x}{(x^2+r^2)^{3/2}} = \frac{\sigma x}{4\epsilon_0}\int_0^R \frac{2r\,dr}{(x^2+r^2)^{3/2}}$$

You can evaluate this integral by making the substitution $t = x^2 + r^2$ (which yields $dt = 2r\,dr$); you can work out the details. The result is

$$E_x = \frac{\sigma x}{2\epsilon_0}\left[-\frac{1}{\sqrt{x^2+R^2}} + \frac{1}{x}\right]$$

$$= \frac{\sigma}{2\epsilon_0}\left[1 - \frac{1}{\sqrt{(R^2/x^2)+1}}\right] \qquad (21.11)$$

EVALUATE If the disk is very large (or if we are very close to it), so that $R \gg x$, the term $1/\sqrt{(R^2/x^2)+1}$ in Eq. (21.11) is very much less than 1. Then Eq. (21.11) becomes

$$E = \frac{\sigma}{2\epsilon_0} \qquad (21.12)$$

Our final result does not contain the distance x from the plane. Hence the electric field produced by an *infinite* plane sheet of charge is *independent of the distance from the sheet.* The field direction is everywhere perpendicular to the sheet, away from it. There is no such thing as an infinite sheet of charge, but if the dimensions of the sheet are much larger than the distance x of the field point P from the sheet, the field is very nearly given by Eq. (21.12).

If P is to the *left* of the plane $(x < 0)$, the result is the same except that the direction of $\vec{E}$ is to the left instead of the right. If the surface charge density is negative, the directions of the fields on both sides of the plane are toward it rather than away from it.

KEYCONCEPT To find the electric field due to a two-dimensional charge distribution such as a disk, divide the distribution into infinitesimal segments such as rings for which you know the components of the electric field. Then integrate over these segments to find the net field.

EXAMPLE 21.12 Field of two oppositely charged infinite sheets

Two infinite plane sheets with uniform surface charge densities $+\sigma$ and $-\sigma$ are placed parallel to each other with separation d (**Fig. 21.26**). Find the electric field between the sheets, above the upper sheet, and below the lower sheet.

IDENTIFY and SET UP Equation (21.12) gives the electric field due to a single infinite plane sheet of charge. To find the field due to *two* such sheets, we combine the fields by using the principle of superposition (Fig. 21.26).

EXECUTE From Eq. (21.12), both $\vec{E}_1$ and $\vec{E}_2$ have the same magnitude at all points, independent of distance from either sheet:

$$E_1 = E_2 = \frac{\sigma}{2\epsilon_0}$$

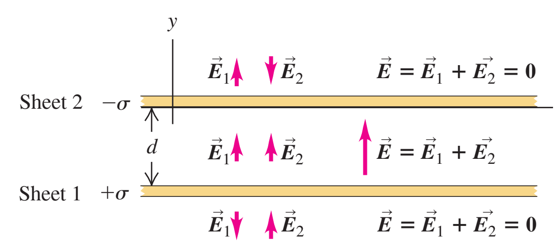
From Example 21.11, $\vec{E}_1$ is everywhere directed away from sheet 1, and $\vec{E}_2$ is everywhere directed toward sheet 2.

Between the sheets, $\vec{E}_1$ and $\vec{E}_2$ reinforce each other; above the upper sheet and below the lower sheet, they cancel each other. Thus the total field is

$$\vec{E} = \vec{E}_1 + \vec{E}_2 = \begin{cases} \mathbf{0} & \text{above the upper sheet} \\ \dfrac{\sigma}{\epsilon_0}\hat{j} & \text{between the sheets} \\ \mathbf{0} & \text{below the lower sheet} \end{cases}$$

EVALUATE Because we considered the sheets to be infinite, our result does not depend on the separation d. Our result shows that the field between oppositely charged plates is essentially uniform if the plate separation is much smaller than the dimensions of the plates. We actually used this result in Example 21.7 (Section 21.4).

Figure **21.26** Finding the electric field due to two oppositely charged infinite sheets. The sheets are seen edge-on; only a portion of the infinite sheets can be shown!



CAUTION **Electric fields are not "flows"** You may have thought that the field $\vec{E}_1$ of sheet 1 would be unable to "penetrate" sheet 2, and that field $\vec{E}_2$ caused by sheet 2 would be unable to "penetrate" sheet 1. You might conclude this if you think of the electric field as some kind of physical substance that "flows" into or out of charges. But there is no such substance, and the electric fields $\vec{E}_1$ and $\vec{E}_2$ depend on only the individual charge distributions that create them. The *total* field at every point is just the vector sum of $\vec{E}_1$ and $\vec{E}_2$.

KEYCONCEPT An infinite, uniform sheet of charge produces a uniform electric field at all points. If the charge is positive, the field points away from the sheet; if the charge is negative, the field points toward the sheet.

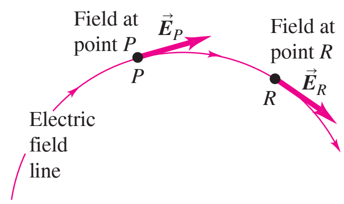
TEST YOUR UNDERSTANDING OF SECTION 21.5 Suppose that the line of charge in Fig. 21.24 (Example 21.10) had charge $+Q$ distributed uniformly between $y = 0$ and $y = +a$ and had charge $-Q$ distributed uniformly between $y = 0$ and $y = -a$. In this situation, the electric field at P would be (i) in the positive x-direction; (ii) in the negative x-direction; (iii) in the positive y-direction; (iv) in the negative y-direction; (v) zero; (vi) none of these.

ANSWER

(iv) Think of a pair of segments of length dy, one at coordinate $y > 0$ and the other at coordinate $-y < 0$. The upper segment has a positive charge and produces an electric field $d\vec{E}$ at P that points away from the segment, so this $d\vec{E}$ has a positive x-component and a negative y-component, like the vector $d\vec{E}$ in Fig. 21.24. The lower segment has the same amount of negative charge. It produces a $d\vec{E}$ that has the same magnitude but points *toward* the lower segment, so it has a negative x-component and a negative y-component. By symmetry, the two x-components are equal but opposite, so they cancel. Thus the total electric field has only a negative y-component.

21.6 ELECTRIC FIELD LINES

Figure 21.27 The direction of the electric field at any point is tangent to the field line through that point.



The concept of an electric field can be a little elusive because you can't see an electric field directly. Electric field *lines* can be a big help for visualizing electric fields and making them seem more real. An **electric field line** is an imaginary line or curve drawn through a region of space so that its tangent at any point is in the direction of the electric-field vector at that point. **Figure 21.27** shows the basic idea. (We used a similar concept in our discussion of fluid flow in Section 12.5. A *streamline* is a line or curve whose tangent at any point is in the direction of the velocity of the fluid at that point. However, the similarity between electric field lines and fluid streamlines is a mathematical one only; there is nothing "flowing" in an electric field.) The English scientist Michael Faraday (1791–1867) first introduced the concept of field lines. He called them "lines of force," but the term "field lines" is preferable.

Electric field lines show the direction of $\vec{E}$ at each point, and their spacing gives a general idea of the *magnitude* of $\vec{E}$ at each point. Where $\vec{E}$ is strong, we draw lines close together; where $\vec{E}$ is weaker, they are farther apart. At any particular point, the electric field has a unique direction, so only one field line can pass through each point of the field. In other words, *field lines never intersect.*

Figure 21.28 shows some of the electric field lines in a plane containing (a) a single positive charge; (b) two equal-magnitude charges, one positive and one negative (a dipole); and (c) two equal positive charges. Such diagrams are called *field maps;* they are cross sections of the actual three-dimensional patterns. The direction of the total electric field at every point in each diagram is along the tangent to the electric field line passing through the point. Arrowheads indicate the direction of the $\vec{E}$-field vector along each field line. The actual field vectors have been drawn at several points in each pattern. Notice that in general, the magnitude of the electric field is different at different points on a given field line; a field line is *not* a curve of constant electric-field magnitude!

Figure 21.28 shows that field lines are directed *away* from positive charges (since close to a positive point charge, $\vec{E}$ points away from the charge) and *toward* negative charges (since close to a negative point charge, $\vec{E}$ points toward the charge). In regions where the field magnitude is large, such as between the positive and negative charges in Fig. 21.28b, the field lines are drawn close together. In regions where the field magnitude is small, such as between the two positive charges in Fig. 21.28c, the lines are widely separated. In a *uniform* field, the field lines are straight, parallel, and uniformly spaced, as in Fig. 21.20.

Figure 21.28 Electric field lines for three different charge distributions. In general, the magnitude of $\vec{E}$ is different at different points along a given field line.

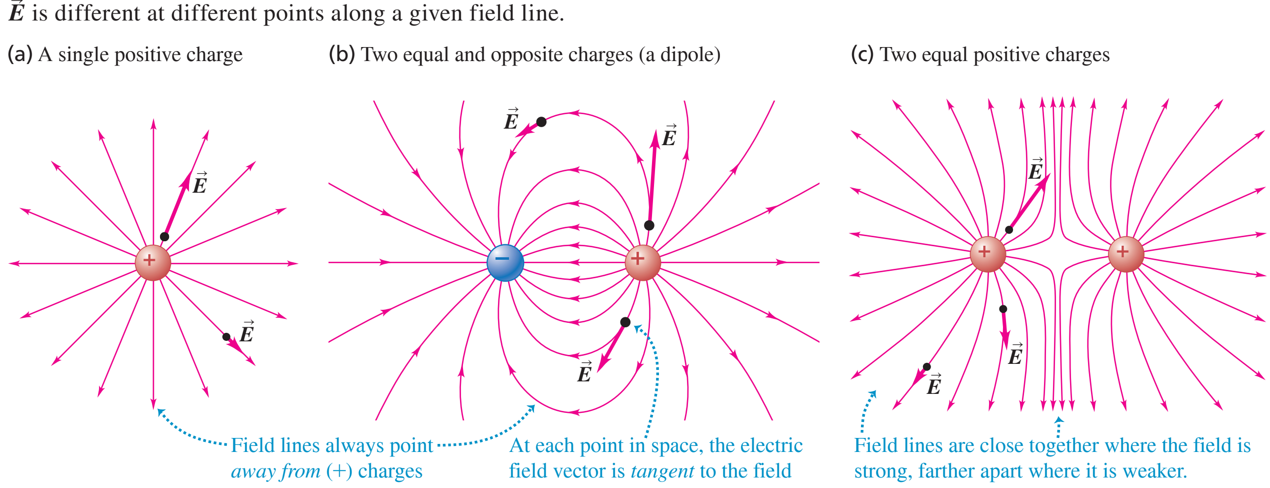
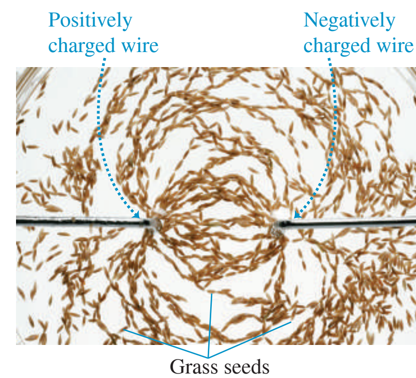


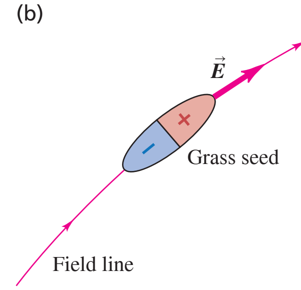
Figure 21.29a is a view from above of a demonstration setup for visualizing electric field lines. In the arrangement shown here, the tips of two positively charged wires are inserted in a container of insulating liquid, and some grass seeds are floated on the liquid. The grass seeds are electrically neutral insulators, but the electric field of the two charged wires causes *polarization* of the grass seeds; there is a slight shifting of the positive and negative charges within the molecules of each seed, like that shown in Fig. 21.8. The positively charged end of each grass seed is pulled in the direction of $\vec{E}$ and the negatively charged end is pulled opposite $\vec{E}$. Hence the long axis of each grass seed tends to orient parallel to the electric field, in the direction of the field line that passes through the position of the seed (Fig. 21.29b).

Figure 21.29 (a) Electric field lines produced by two opposite charges. The pattern is formed by grass seeds floating in mineral oil; the charges are on two wires whose tips are inserted into the oil. Compare this pattern with Fig. 21.28b. (b) The electric field causes polarization of the grass seeds, which in turn causes the seeds to align with the field.

(a)



(b)



CAUTION **Electric field lines are not trajectories** It's a common misconception that if a particle of charge q is in motion where there is an electric field, the particle must move along an electric field line. Because $\vec{E}$ at any point is tangent to the field line that passes through that point, it is true that the *force* $\vec{F} = q\vec{E}$ on the particle, and hence the particle's acceleration, are tangent to the field line. But we learned in Chapter 3 that when a particle moves on a curved path, its acceleration *cannot* be tangent to the path. In general, the trajectory of a charged particle is *not* the same as a field line.

TEST YOUR UNDERSTANDING OF SECTION 21.6 Suppose the electric field lines in a region of space are straight lines. If a charged particle is released from rest in that region, will the trajectory of the particle be along a field line?

ANSWER

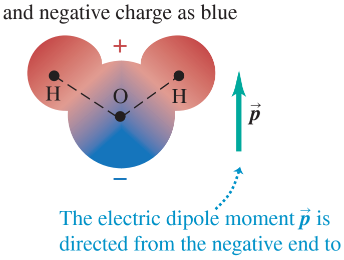
yes If the field lines are straight, $\vec{E}$ must point in the same direction throughout the region. Hence the force $\vec{F} = q\vec{E}$ on a particle of charge q is always in the same direction. A particle released from rest accelerates in a straight line in the direction of $\vec{F}$, and so its trajectory is a straight line along a field line.

21.7 ELECTRIC DIPOLES

An **electric dipole** is a pair of point charges with equal magnitude and opposite sign (a positive charge q and a negative charge $-q$) separated by a distance d. We introduced electric dipoles in Example 21.8 (Section 21.5); the concept is worth exploring further because many physical systems, from molecules to TV antennas, can be described as electric dipoles. We'll also use this concept extensively in our discussion of dielectrics in Chapter 24.

Figure 21.30a shows a molecule of water (H_2O), which in many ways behaves like an electric dipole. The water molecule as a whole is electrically neutral, but the chemical bonds within the molecule cause a displacement of charge; the result is a net negative charge on the oxygen end of the molecule and a net positive charge on the hydrogen end, forming an electric dipole. The effect is equivalent to shifting one electron only about 4×10^{-11} m (about the radius of a hydrogen atom), but the consequences of this shift are profound. Water is an excellent solvent for ionic substances such as table salt (sodium chloride, NaCl) precisely because the water molecule is an electric dipole (Fig. 21.30b). When dissolved in water, salt dissociates into a positive sodium ion (Na^+)

(a) A water molecule, showing positive charge as red and negative charge as blue



(b) Various substances dissolved in water

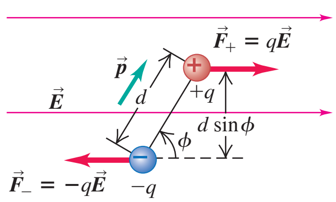
Figure 21.30 (a) A water molecule is an example of an electric dipole. (b) Each test tube contains a solution of a different substance in water. The large electric dipole moment of water makes it an excellent solvent.

and a negative chlorine ion (Cl^-), which tend to be attracted to the negative and positive ends, respectively, of water molecules; this holds the ions in solution. If water molecules were not electric dipoles, water would be a poor solvent, and almost all of the chemistry that occurs in aqueous solutions would be impossible. This includes all of the biochemical reactions that occur in all of the life on earth. In a very real sense, your existence as a living being depends on electric dipoles!

We examine two questions about electric dipoles. First, what forces and torques does an electric dipole experience when placed in an external electric field (that is, a field set up by charges outside the dipole)? Second, what electric field does an electric dipole itself produce?

Force and Torque on an Electric Dipole

Figure **21.31** The net force on this electric dipole is zero, but there is a torque directed into the page that tends to rotate the dipole clockwise.



To start with the first question, let's place an electric dipole in a *uniform* external electric field $\vec{E}$, as shown in **Fig. 21.31**. Both forces $\vec{F}_+$ and $\vec{F}_-$ on the two charges have magnitude qE, but their directions are opposite, and they add to zero. *The net force on an electric dipole in a uniform external electric field is zero.*

However, the two forces don't act along the same line, so their *torques* don't add to zero. We calculate torques with respect to the center of the dipole. Let the angle between the electric field $\vec{E}$ and the dipole axis be ϕ; then the lever arm for both $\vec{F}_+$ and $\vec{F}_-$ is $(d/2)\sin\phi$. The torque of $\vec{F}_+$ and the torque of $\vec{F}_-$ both have the same magnitude of $(qE)(d/2)\sin\phi$, and both torques tend to rotate the dipole clockwise (that is, $\vec{\tau}$ is directed into the page in Fig. 21.31). Hence the magnitude of the net torque is twice the magnitude of either individual torque:

$$\tau = (qE)(d\sin\phi) \qquad (21.13)$$

where $d\sin\phi$ is the perpendicular distance between the lines of action of the two forces.

The product of the charge q and the separation d is the magnitude of a quantity called the **electric dipole moment,** denoted by p:

$$p = qd \qquad \text{(magnitude of electric dipole moment)} \qquad (21.14)$$

The units of p are charge times distance $(\mathrm{C \cdot m})$. For example, the magnitude of the electric dipole moment of a water molecule is $p = 6.13 \times 10^{-30}\ \mathrm{C \cdot m}$.

CAUTION **The symbol p has multiple meanings** Do not confuse dipole moment with momentum or pressure. There aren't as many letters in the alphabet as there are physical quantities, so some letters are used several times. The context usually makes it clear what we mean, but be careful.

We further define the electric dipole moment to be a *vector* quantity $\vec{p}$. The magnitude of $\vec{p}$ is given by Eq. (21.14), and its direction is along the dipole axis from the negative charge to the positive charge as shown in Fig. 21.31.

In terms of p, Eq. (21.13) for the magnitude τ of the torque exerted by the field becomes

Magnitude of electric field $\vec{E}$

Magnitude of torque on an electric dipole

$$\tau = pE\sin\phi \qquad (21.15)$$

Angle between $\vec{p}$ and $\vec{E}$

Magnitude of electric dipole moment $\vec{p}$

Since the angle ϕ in Fig. 21.31 is the angle between the directions of the vectors $\vec{p}$ and $\vec{E}$, this is reminiscent of the expression for the magnitude of the *vector product* discussed in Section 1.10. (You may want to review that discussion.) Hence we can write the torque on the dipole in vector form as

Vector torque on an electric dipole

$$\vec{\tau} = \vec{p} \times \vec{E} \tag{21.16}$$

$\vec{p}$: Electric dipole moment; $\vec{E}$: Electric field

You can use the right-hand rule for the vector product to verify that in the situation shown in Fig. 21.31, $\vec{\tau}$ is directed into the page. The torque is greatest when $\vec{p}$ and $\vec{E}$ are perpendicular and is zero when they are parallel or antiparallel. The torque always tends to turn $\vec{p}$ to line it up with $\vec{E}$. The position $\phi = 0$, with $\vec{p}$ parallel to $\vec{E}$, is a position of stable equilibrium, and the position $\phi = \pi$, with $\vec{p}$ and $\vec{E}$ antiparallel, is a position of unstable equilibrium. The polarization of a grass seed in the apparatus of Fig. 21.29b gives it an electric dipole moment; the torque exerted by $\vec{E}$ then causes the seed to align with $\vec{E}$ and hence with the field lines.

Potential Energy of an Electric Dipole

When a dipole changes direction in an electric field, the electric-field torque does *work* on it, with a corresponding change in potential energy. The work dW done by a torque τ during an infinitesimal displacement $d\phi$ is given by Eq. (10.19): $dW = \tau\, d\phi$. Because the torque is in the direction of decreasing ϕ, we must write the torque as $\tau = -pE\sin\phi$, and

$$dW = \tau\, d\phi = -pE\sin\phi\, d\phi$$

In a finite displacement from ϕ_1 to ϕ_2 the total work done on the dipole is

$$\begin{aligned} W &= \int_{\phi_1}^{\phi_2} (-pE\sin\phi)\, d\phi \\ &= pE\cos\phi_2 - pE\cos\phi_1 \end{aligned}$$

The work is the negative of the change of potential energy, just as in Chapter 7: $W = U_1 - U_2$. So a suitable definition of potential energy U for this system is

$$U(\phi) = -pE\cos\phi \tag{21.17}$$

In this expression we recognize the *scalar product* $\vec{p} \cdot \vec{E} = pE\cos\phi$, so we can also write

Potential energy for an electric dipole in an electric field

$$U = -\vec{p} \cdot \vec{E} \tag{21.18}$$

$\vec{E}$: Electric field; $\vec{p}$: Electric dipole moment

The potential energy has its minimum (most negative) value $U = -pE$ at the stable equilibrium position, where $\phi = 0$ and $\vec{p}$ is parallel to $\vec{E}$. The potential energy is maximum when $\phi = \pi$ and $\vec{p}$ is antiparallel to $\vec{E}$; then $U = +pE$. At $\phi = \pi/2$, where $\vec{p}$ is perpendicular to $\vec{E}$, U is zero. We could define U differently so that it is zero at some other orientation of $\vec{p}$, but our definition is simplest.

Equation (21.18) gives us another way to look at the effect illustrated in Fig. 21.29. The electric field $\vec{E}$ gives each grass seed an electric dipole moment, and the grass seed then aligns itself with $\vec{E}$ to minimize the potential energy.

EXAMPLE 21.13 Force and torque on an electric dipole

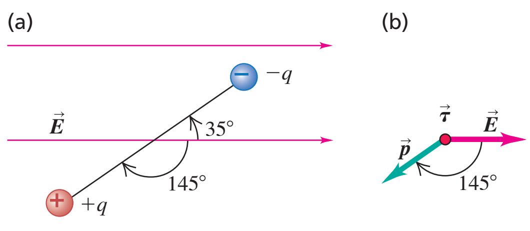
WITH VARIATION PROBLEMS

Figure 21.32a (next page) shows an electric dipole in a uniform electric field of magnitude 5.0×10^5 N/C that is directed parallel to the plane of the figure. The charges are $\pm 1.6 \times 10^{-19}$ C; both lie in the plane and are separated by 0.125 nm $= 0.125 \times 10^{-9}$ m. Find (a) the net force exerted by the field on the dipole; (b) the magnitude and direction of the electric dipole moment; (c) the magnitude and direction of the torque; (d) the potential energy of the system in the position shown.

IDENTIFY and SET UP This problem uses the ideas of this section about an electric dipole placed in an electric field. We use the relationship $\vec{F} = q\vec{E}$ for each point charge to find the force on the dipole as a whole. Equation (21.14) gives the dipole moment, Eq. (21.16) gives the torque on the dipole, and Eq. (21.18) gives the potential energy of the system.

EXECUTE (a) The field is uniform, so the forces on the two charges are equal and opposite. Hence the total force on the dipole is zero.

Figure **21.32** (a) An electric dipole. (b) Directions of the electric dipole moment, electric field, and torque ($\vec{\tau}$ points out of the page).



(b) The magnitude p of the electric dipole moment $\vec{p}$ is

$$p = qd = (1.6 \times 10^{-19}\ \text{C})(0.125 \times 10^{-9}\ \text{m}) = 2.0 \times 10^{-29}\ \text{C}\cdot\text{m}$$

The direction of $\vec{p}$ is from the negative to the positive charge, 145° clockwise from the electric-field direction (Fig. 21.32b).

(c) The magnitude of the torque is

$$\tau = pE\sin\phi = (2.0 \times 10^{-29}\ \text{C}\cdot\text{m})(5.0 \times 10^5\ \text{N/C})(\sin 145°) = 5.7 \times 10^{-24}\ \text{N}\cdot\text{m}$$

From the right-hand rule for vector products (see Section 1.10), the direction of the torque $\vec{\tau} = \vec{p} \times \vec{E}$ is out of the page. This corresponds to a counterclockwise torque that tends to align $\vec{p}$ with $\vec{E}$.

(d) The potential energy

$$\begin{aligned} U &= -pE\cos\phi \\ &= -(2.0 \times 10^{-29}\ \text{C}\cdot\text{m})(5.0 \times 10^5\ \text{N/C})(\cos 145°) \\ &= 8.2 \times 10^{-24}\ \text{J} \end{aligned}$$

EVALUATE The charge magnitude, the distance between the charges, the dipole moment, and the potential energy are all very small, but are all typical of molecules.

KEYCONCEPT An electric dipole placed in an electric field tends to orient itself with its dipole moment in the same direction as the field.

BIO APPLICATION A Fish with an Electric Dipole Moment Unlike the tiger shark (Section 21.4), which senses the electric fields produced by its prey, the African knifefish *Gymnarchus niloticus*—which is nocturnal and has poor vision—hunts other fish by generating its own electric field. It can make its tail negatively charged relative to its head, producing a dipole-like field similar to that in Fig. 21.28b. When a smaller fish ventures into the field, its body alters the field pattern and alerts *G. niloticus* that a meal is present.

In this discussion we have assumed that $\vec{E}$ is uniform, so there is no net force on the dipole. If $\vec{E}$ is not uniform, the forces at the ends may not cancel completely, and the net force may not be zero. Thus an object with zero net charge but an electric dipole moment can experience a net force in a nonuniform electric field. As we mentioned in Section 21.1, an uncharged object can be polarized by an electric field, giving rise to a separation of charge and an electric dipole moment. This is how uncharged objects can experience electrostatic forces (see Fig. 21.8).

Field of an Electric Dipole

Now let's think of an electric dipole as a *source* of electric field. Figure 21.28b shows the general shape of the field due to a dipole. At each point in the pattern the total $\vec{E}$ field is the vector sum of the fields from the two individual charges, as in Example 21.8 (Section 21.5). Try drawing diagrams showing this vector sum for several points.

To get quantitative information about the field of an electric dipole, we have to do some calculating, as illustrated in the next example. Notice the use of the principle of superposition of electric fields to add up the contributions to the field of the individual charges. Also notice that we need to use approximation techniques even for the relatively simple case of a field due to two charges. Field calculations often become very complicated, and computer analysis is typically used to determine the field due to an arbitrary charge distribution.

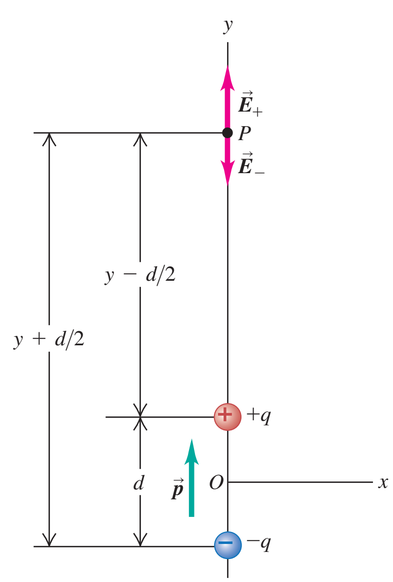
EXAMPLE 21.14 Field of an electric dipole, revisited

WITH VARIATION PROBLEMS

An electric dipole is centered at the origin, with $\vec{p}$ in the direction of the $+y$-axis (**Fig. 21.33**). Derive an approximate expression for the electric field at a point P on the y-axis for which y is much larger than d. To do this, use the binomial expansion $(1 + x)^n \cong 1 + nx + n(n-1)x^2/2 + \cdots$ (valid for the case $|x| < 1$).

IDENTIFY and SET UP We use the principle of superposition: The total electric field is the vector sum of the field produced by the positive charge and the field produced by the negative charge. At the field point P shown in Fig. 21.33, the field $\vec{E}_+$ of the positive charge has a positive (upward) y-component and the field $\vec{E}_-$ of the negative charge has a negative (downward) y-component. We add these components to find the total field and then apply the approximation that y is much greater than d.

Figure **21.33** Finding the electric field of an electric dipole at a point on its axis.



EXECUTE The total y-component E_y of electric field from the two charges is

$$E_y = \frac{q}{4\pi\epsilon_0}\left[\frac{1}{(y - d/2)^2} - \frac{1}{(y + d/2)^2}\right]$$

$$= \frac{q}{4\pi\epsilon_0 y^2}\left[\left(1 - \frac{d}{2y}\right)^{-2} - \left(1 + \frac{d}{2y}\right)^{-2}\right]$$

We used this same approach in Example 21.8 (Section 21.5). Now the approximation: When we are far from the dipole compared to its size, so $y \gg d$, we have $d/2y \ll 1$. With $n = -2$ and with $d/2y$ replacing x in the binomial expansion, we keep only the first two terms (the terms we discard are much smaller). We then have

$$\left(1 - \frac{d}{2y}\right)^{-2} \cong 1 + \frac{d}{y} \quad \text{and} \quad \left(1 + \frac{d}{2y}\right)^{-2} \cong 1 - \frac{d}{y}$$

Hence E_y is given approximately by

$$E_y \cong \frac{q}{4\pi\epsilon_0 y^2}\left[1 + \frac{d}{y} - \left(1 - \frac{d}{y}\right)\right] = \frac{qd}{2\pi\epsilon_0 y^3} = \frac{p}{2\pi\epsilon_0 y^3}$$

EVALUATE An alternative route to this result is to put the fractions in the first expression for E_y over a common denominator, add, and then approximate the denominator $(y - d/2)^2(y + d/2)^2$ as y^4. We leave the details to you.

For points P off the coordinate axes, the expressions are more complicated, but at *all* points far away from the dipole (in any direction) the field drops off as $1/r^3$. We can compare this with the $1/r^2$ behavior of a point charge, the $1/r$ behavior of a long line charge, and the independence of r for a large sheet of charge. There are charge distributions for which the field drops off even more quickly. At large distances, the field of an *electric quadrupole,* which consists of two equal dipoles with opposite orientation, separated by a small distance, drops off as $1/r^4$.

KEYCONCEPT The electric field produced by an electric dipole is the vector sum of the fields produced by the positive and negative charges that make up the dipole. At distances r from the dipole that are large compared to the dipole size, the field magnitude is inversely proportional to the cube of r.

TEST YOUR UNDERSTANDING OF SECTION 21.7 An electric dipole is placed in a region of uniform electric field $\vec{E}$, with the electric dipole moment $\vec{p}$ pointing in the direction opposite to $\vec{E}$. Is the dipole (i) in stable equilibrium, (ii) in unstable equilibrium, or (iii) neither? (*Hint:* You many want to review Section 7.5.)

ANSWER

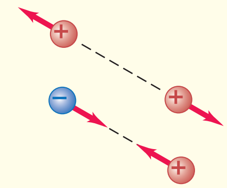
❙ **(ii)** Equations (21.17) and (21.18) tell us that the potential energy for a dipole in an electric field is $U = -\vec{p}\cdot\vec{E} = -pE\cos\phi$, where ϕ is the angle between the directions of $\vec{p}$ and $\vec{E}$. If $\vec{p}$ and $\vec{E}$ point in opposite directions, so that $\phi = 180°$, we have $\cos\phi = -1$ and $U = +pE$. This is the maximum value that U can have. From our discussion of energy diagrams in Section 7.5, it follows that this is a situation of unstable equilibrium.

CHAPTER 21 SUMMARY

Electric charge, conductors, and insulators: The fundamental quantity in electrostatics is electric charge. There are two kinds of charge, positive and negative. Charges of the same sign repel each other; charges of opposite sign attract. Charge is conserved; the total charge in an isolated system is constant.

All ordinary matter is made of protons, neutrons, and electrons. The positive protons and electrically neutral neutrons in the nucleus of an atom are bound together by the nuclear force; the negative electrons surround the nucleus at distances much greater than the nuclear size. Electric interactions are chiefly responsible for the structure of atoms, molecules, and solids.

Conductors are materials in which charge moves easily; in insulators, charge does not move easily. Most metals are good conductors; most nonmetals are insulators.

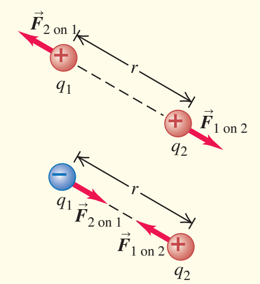


Coulomb's law: For charges q_1 and q_2 separated by a distance r, the magnitude of the electric force on either charge is proportional to the product $q_1 q_2$ and inversely proportional to r^2. The force on each charge is along the line joining the two charges—repulsive if q_1 and q_2 have the same sign, attractive if they have opposite signs. In SI units the unit of electric charge is the coulomb, abbreviated C. (See Examples 21.1 and 21.2.)

When two or more charges each exert a force on a charge, the total force on that charge is the vector sum of the forces exerted by the individual charges. (See Examples 21.3 and 21.4.)

$$F = \frac{1}{4\pi\epsilon_0}\frac{|q_1 q_2|}{r^2} \qquad (21.2)$$

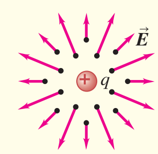
$$\frac{1}{4\pi\epsilon_0} = 8.988 \times 10^9\ \mathrm{N \cdot m^2/C^2}$$



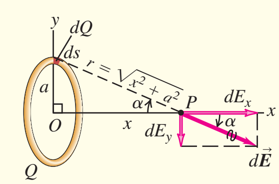
Electric field: Electric field $\vec{E}$, a vector quantity, is the force per unit charge exerted on a test charge at any point. The electric field produced by a point charge is directed radially away from a positive charge and radially toward a negative charge. (See Examples 21.5–21.7.)

$$\vec{E} = \frac{\vec{F}_0}{q_0} \qquad (21.3)$$

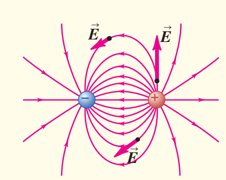
$$\vec{E} = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}\hat{r} \qquad (21.7)$$



Superposition of electric fields: The electric field $\vec{E}$ of any combination of charges is the vector sum of the fields caused by the individual charges. To calculate the electric field caused by a continuous distribution of charge, divide the distribution into small elements, calculate the field caused by each element, and then carry out the vector sum, usually by integrating. Charge distributions are described by linear charge density λ, surface charge density σ, and volume charge density ρ. (See Examples 21.8–21.12.)



Electric field lines: Field lines provide a graphical representation of electric fields. At any point on a field line, the tangent to the line is in the direction of $\vec{E}$ at that point. The number of lines per unit area (perpendicular to their direction) is proportional to the magnitude of $\vec{E}$ at the point.

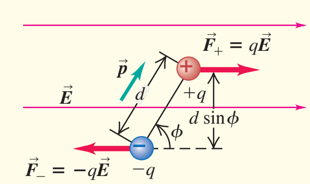


Electric dipoles: An electric dipole is a pair of electric charges of equal magnitude q but opposite sign, separated by a distance d. The electric dipole moment $\vec{p}$ has magnitude $p = qd$. The direction of $\vec{p}$ is from negative toward positive charge. An electric dipole in an electric field $\vec{E}$ experiences a torque $\vec{\tau}$ equal to the vector product of $\vec{p}$ and $\vec{E}$. The magnitude of the torque depends on the angle ϕ between $\vec{p}$ and $\vec{E}$. The potential energy U for an electric dipole in an electric field also depends on the relative orientation of $\vec{p}$ and $\vec{E}$. (See Examples 21.13 and 21.14.)

$$\tau = pE\sin\phi \qquad (21.15)$$

$$\vec{\tau} = \vec{p} \times \vec{E} \qquad (21.16)$$

$$U = -\vec{p} \cdot \vec{E} \qquad (21.18)$$



Chapter 21 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 21.3 and 21.4 (Section 21.3) before attempting these problems.

VP21.4.1 Charge $q_1 = +4.00$ nC is on the x-axis at $x = 0$ and charge $q_2 = -1.20$ nC is on the x-axis at $x = 6.00$ cm. Find the total electric force exerted by q_1 and q_2 on a third charge $q_3 = -2.00$ nC on the x-axis at $x = 4.00$ cm.

VP21.4.2 Charge $q_1 = +3.60$ nC is on the x-axis at $x = 0$ and an unknown charge q_2 is on the x-axis at $x = -4.00$ cm. The total electric force that q_1 and q_2 exert on a third charge $q_3 = +2.00$ nC on the x-axis at $x = 4.00$ cm is in the $+x$-direction and has magnitude 54.0 μN. Find (a) the magnitude and direction of the force that q_1 exerts on q_3, (b) the magnitude and direction of the force that q_2 exerts on q_3, and (c) the charge q_2.

VP21.4.3 Charge $q_1 = +6.00$ nC is at $x = 0$, $y = 0.200$ m and charge $q_2 = -6.00$ nC is at $x = 0$, $y = -0.200$ m. Find the magnitude and direction of the total electric force exerted by q_1 and q_2 on a third charge $q_3 = -2.50$ nC at $x = 0.150$ m, $y = 0$.

VP21.4.4 A small piece of plastic with charge $q_1 = +4.00$ nC is at $x = 0$, $y = 0.200$ m and a second small piece of plastic with charge $q_2 = -4.00$ nC is at $x = 0$, $y = 0$. Find the x- and y-components of the total electric force exerted by q_1 and q_2 on a third small piece of plastic with charge $q_3 = -1.50$ nC at $x = 0.250$ m, $y = 0$.

Be sure to review EXAMPLES 21.8, 21.9, and 21.10 (Section 21.5) before attempting these problems.

VP21.10.1 Charge $q_1 = +4.00$ nC is at $x = 0$, $y = 0.200$ m and charge $q_2 = -5.00$ nC is at $x = 0$, $y = 0$. Find the x- and y-components of the total electric field caused by q_1 and q_2 at (a) $x = 0$, $y = 0.100$ m; (b) $x = 0$, $y = 0.400$ m; (c) $x = 0.200$ m, $y = 0$.

VP21.10.2 Charge $q_1 = +1.80$ nC is on the x-axis at $x = 0$ and an unknown charge q_2 is on the x-axis at $x = +4.00$ cm. The total electric field caused by q_1 and q_2 at a point P on the x-axis at $x = +2.00$ cm is in the $+x$-direction and has magnitude 6.75×10^4 N/C. Find (a) the magnitude and direction of the electric field that q_1 causes at P, (b) the magnitude and direction of the electric field that q_2 causes at P, and (c) the charge q_2.

VP21.10.3 A simple model of a hydrogen atom is a positive point charge $+e$ (representing the proton) at the center of a ring of radius a with negative charge $-e$ distributed uniformly around the ring (representing the electron in orbit around the proton). (a) Find the magnitude of the total electric field due to this charge distribution at a point a distance a from the proton and perpendicular to the plane of the ring. (b) Does the field point toward or away from the proton?

VP21.10.4 Positive charge Q is distributed uniformly along a rod of length L that lies along the x-axis from $x = L$ to $x = 2L$. Find (a) how much charge is contained within a segment of the rod of length dx, (b) the x- and y-components of the infinitesimal electric field caused at the origin ($x = 0$) by one such segment at coordinate x, and (c) the x- and y-components of the total electric field caused at the origin by the rod.

Be sure to review EXAMPLES 21.13 and 21.14 (Section 21.7) before attempting these problems.

VP21.14.1 The electric dipole moment of a water molecule (H_2O) is 6.13×10^{-30} C · m. You place a water molecule in a uniform electric field of magnitude 3.00×10^5 N/C, with the dipole moment vector at an angle of 50.0° from the direction of the field. Calculate (a) the magnitude of the torque on the water molecule and (b) the potential energy for the water molecule.

VP21.14.2 An electric dipole is in a uniform electric field of magnitude 8.50×10^4 N/C. The charges in the dipole are separated by 1.10×10^{-10} m, and the torque on the dipole when its dipole moment is perpendicular to the electric field is 6.60×10^{-26} N · m. Calculate the magnitude of the charges that make up the dipole.

VP21.14.3 You place a molecule in a uniform electric field of magnitude 1.20×10^5 N/C. It requires 4.60×10^{-25} J of work to change the orientation of the molecule's dipole moment from parallel to the field to opposite to the field. Calculate the magnitude of the dipole moment.

VP21.14.4 The ionic molecule potassium bromide (KBr), made up of a positive potassium ion (K^+) of charge $+e = 1.60 \times 10^{-19}$ C and a negative bromine ion (Br^-) of charge $-e = -1.60 \times 10^{-19}$, has an electric dipole moment of 3.50×10^{-29} C · m. (a) Calculate the distance between the two ions. (b) At a certain point along the axis that connects the ions, the electric field due to the KBr molecule has magnitude 8.00×10^4 N/C. How far from the center of the molecule is this point?

BRIDGING PROBLEM Calculating Electric Field: Half a Ring of Charge

Positive charge Q is uniformly distributed around a semicircle of radius a as shown in **Fig. 21.34**. Find the magnitude and direction of the resulting electric field at point P, the center of curvature of the semicircle.

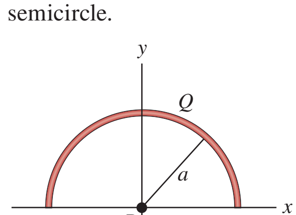


Figure 21.34 Charge uniformly distributed around a semicircle.

SOLUTION GUIDE

IDENTIFY and SET UP

1. The target variables are the components of the electric field at P.
2. Divide the semicircle into infinitesimal segments, each of which is a short circular arc of radius a and angle $d\theta$. What is the length of such a segment? How much charge is on a segment?
3. Consider an infinitesimal segment located at an angular position θ on the semicircle, measured from the lower right corner of the semicircle at $x = a, y = 0$. (Thus $\theta = \pi/2$ at $x = 0, y = a$ and $\theta = \pi$ at $x = -a$, $y = 0$.) What are the x- and y-components of the electric field at P (dE_x and dE_y) produced by just this segment?

EXECUTE

4. Integrate your expressions for dE_x and dE_y from $\theta = 0$ to $\theta = \pi$. The results will be the x-component and y-component of the electric field at P.
5. Use your results from step 4 to find the magnitude and direction of the field at P.

EVALUATE

6. Does your result for the electric-field magnitude have the correct units?
7. Explain how you could have found the x-component of the electric field by using a symmetry argument.
8. What would be the electric field at P if the semicircle were extended to a full circle centered at P?

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q21.1 If you peel two strips of transparent tape off the same roll and immediately let them hang near each other, they will repel each other. If you then stick the sticky side of one to the shiny side of the other and rip them apart, they will attract each other. Give a plausible explanation, involving transfer of electrons between the strips of tape, for this sequence of events.

Q21.2 Two metal spheres are hanging from nylon threads. When you bring the spheres close to each other, they tend to attract. Based on this information alone, discuss all the possible ways that the spheres could be charged. Is it possible that after the spheres touch, they will cling together? Explain.

Q21.3 The electric force between two charged particles becomes weaker with increasing distance. Suppose instead that the electric force were *independent* of distance. In this case, would a charged comb still cause a neutral insulator to become polarized as in Fig. 21.8? Why or why not? Would the neutral insulator still be attracted to the comb? Again, why or why not?

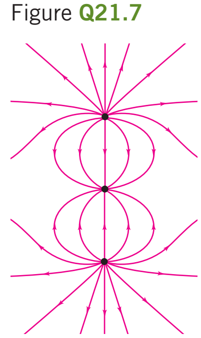
Q21.4 Your clothing tends to cling together after going through the dryer. Why? Would you expect more or less clinging if all your clothing were made of the same material (say, cotton) than if you dried different kinds of clothing together? Again, why? (You may want to experiment with your next load of laundry.)

Q21.5 An uncharged metal sphere hangs from a nylon thread. When a positively charged glass rod is brought close to the metal sphere, the sphere is drawn toward the rod. But if the sphere touches the rod, it suddenly flies away from the rod. Explain why the sphere is first attracted and then repelled.

Q21.6 BIO Estimate how many electrons there are in your body. Make any assumptions you feel are necessary, but clearly state what they are. (*Hint:* Most of the atoms in your body have equal numbers of electrons, protons, and neutrons.) What is the combined charge of all these electrons?

Q21.7 **Figure Q21.7** shows some of the electric field lines due to three point charges arranged along the vertical axis. All three charges have the same magnitude. (a) What are the signs of the three charges? Explain your reasoning. (b) At what point(s) is the magnitude of the electric field the smallest? Explain your reasoning. Explain how the fields produced by each individual point charge combine to give a small net field at this point or points.

Figure **Q21.7**



Q21.8 Good conductors of electricity, such as metals, are typically good conductors of heat; insulators, such as wood, are typically poor conductors of heat. Explain why there is a relationship between conduction of electricity and conduction of heat in these materials.

Q21.9 Suppose that the charge shown in Fig. 21.28a is fixed in position. A small, positively charged particle is then placed at some location and released from rest. Will the trajectory of the particle follow an electric field line? Why or why not? Suppose instead that the particle is placed at some point in Fig. 21.28b and released from rest (the positive and negative charges shown are fixed in position). Will its trajectory follow an electric field line? Again, why or why not? Explain any differences between your answers for the two situations.

Q21.10 Two identical metal objects are mounted on insulating stands. Describe how you could place charges of opposite sign but exactly equal magnitude on the two objects.

Q21.11 Because the charges on the electron and proton have the same absolute value, atoms are electrically neutral. Suppose that this is not precisely true, and the absolute value of the charge of the electron is less than the charge of the proton by 0.00100%. Estimate what the net charge of this textbook would be under these circumstances. Make any assumptions you feel are justified, but state clearly what they are. (*Hint:* Most of the atoms in this textbook have equal numbers of electrons, protons, and neutrons.) What would be the magnitude of the electric force between two textbooks placed 5.0 m apart? Would this force be attractive or repulsive? Discuss how the fact that ordinary matter is stable shows that the absolute values of the charges on the electron and proton must be identical to a *very* high level of accuracy.

Q21.12 If you walk across a nylon rug and then touch a large metal object such as a doorknob, you may get a spark and a shock. Why does this tend to happen more on dry days than on humid days? (*Hint:* See Fig. 21.30.) Why are you less likely to get a shock if you touch a *small* metal object, such as a paper clip?

Q21.13 You have a negatively charged object. How can you use it to place a net negative charge on an insulated metal sphere? To place a net positive charge on the sphere?

Q21.14 When two point charges of equal mass and charge are released on a frictionless table, each has an initial acceleration (magnitude) a_0. If instead you keep one fixed and release the other one, what will be its initial acceleration: a_0, $2a_0$, or $a_0/2$? Explain.

Q21.15 A point charge of mass m and charge Q and another point charge of mass m but charge $2Q$ are released on a frictionless table. If the charge Q has an initial acceleration a_0, what will be the acceleration of $2Q$: a_0, $2a_0$, $4a_0$, $a_0/2$, or $a_0/4$? Explain.

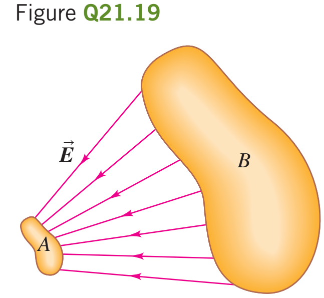
Q21.16 A proton is placed in a uniform electric field and then released. Then an electron is placed at this same point and released. Do these two particles experience the same force? The same acceleration? Do they move in the same direction when released?

Q21.17 In Example 21.1 (Section 21.3) we saw that the electric force between two α particles is of the order of 10^{35} times as strong as the gravitational force. So why do we readily feel the gravity of the earth but no electric force from it?

Q21.18 What similarities do electric forces have with gravitational forces? What are the most significant differences?

Q21.19 Two irregular objects A and B carry charges of opposite sign. **Figure Q21.19** shows the electric field lines near each of these objects. (a) Which object is positive, A or B? How do you know? (b) Where is the electric field stronger, close to A or close to B? How do you know?

Figure **Q21.19**

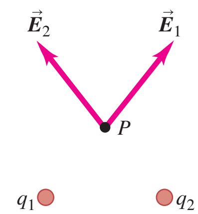


Q21.20 Atomic nuclei are made of protons and neutrons. This shows that there must be another kind of interaction in addition to gravitational and electric forces. Explain.

Q21.21 Sufficiently strong electric fields can cause atoms to become positively ionized—that is, to lose one or more electrons. Explain how this can happen. What determines how strong the field must be to make this happen?

Q21.22 The electric fields at point P due to the positive charges q_1 and q_2 are shown in **Fig. Q21.22**. Does the fact that they cross each other violate the statement in Section 21.6 that electric field lines never cross? Explain.

Figure **Q21.22**



Q21.23 The air temperature and the velocity of the air have different values at different places in the earth's atmosphere. Is the air velocity a vector field? Why or why not? Is the air temperature a vector field? Again, why or why not?

EXERCISES

Section 21.3 Coulomb's Law

21.1 •• Excess electrons are placed on a small lead sphere with mass 8.00 g so that its net charge is -3.20×10^{-9} C. (a) Find the number of excess electrons on the sphere. (b) How many excess electrons are there per lead atom? The atomic number of lead is 82, and its atomic mass is 207 g/mol.

21.2 • Lightning occurs when there is a flow of electric charge (principally electrons) between the ground and a thundercloud. The maximum rate of charge flow in a lightning bolt is about 20,000 C/s; this lasts for 100 μs or less. How much charge flows between the ground and the cloud in this time? How many electrons flow during this time?

21.3 •• If a proton and an electron are released when they are 2.0×10^{-10} m apart (a typical atomic distance), find the initial acceleration of each particle.

21.4 •• In Example 21.4, what is the net force (magnitude and direction) on charge q_1 exerted by the other two charges?

21.5 • BIO **Signal Propagation in Neurons.** *Neurons* are components of the nervous system of the body that transmit signals as electric impulses travel along their length. These impulses propagate when charge suddenly rushes into and then out of a part of the neuron called an *axon*. Measurements have shown that, during the inflow part of this cycle, approximately 5.6×10^{11} Na^+ (sodium ions) per meter, each with charge $+e$, enter the axon. How many coulombs of charge enter a 1.5 cm length of the axon during this process?

21.6 • Two small spheres spaced 20.0 cm apart have equal charge. How many excess electrons must be present on each sphere if the magnitude of the force of repulsion between them is 3.33×10^{-21} N?

21.7 •• An average human weighs about 650 N. If each of two average humans could carry 1.0 C of excess charge, one positive and one negative, how far apart would they have to be for the electric attraction between them to equal their 650 N weight?

21.8 •• Two small aluminum spheres, each having mass 0.0250 kg, are separated by 80.0 cm. (a) How many electrons does each sphere contain? (The atomic mass of aluminum is 26.982 g/mol, and its atomic number is 13.) (b) How many electrons would have to be removed from one sphere and added to the other to cause an attractive force between the spheres of magnitude 1.00×10^4 N (roughly 1 ton)? Assume that the spheres may be treated as point charges. (c) What fraction of all the electrons in each sphere does this represent?

21.9 •• Two small plastic spheres are given positive electric charges. When they are 15.0 cm apart, the repulsive force between them has magnitude 0.220 N. What is the charge on each sphere (a) if the two charges are equal and (b) if one sphere has four times the charge of the other?

21.10 •• Two point charges are placed on the x-axis as follows: Charge $q_1 = +4.00$ nC is located at $x = 0.200$ m, and charge $q_2 = +5.00$ nC is at $x = -0.300$ m. What are the magnitude and direction of the total force exerted by these two charges on a negative point charge $q_3 = -6.00$ nC that is placed at the origin?

21.11 • In an experiment in space, one proton is held fixed and another proton is released from rest a distance of 2.50 mm away. (a) What is the initial acceleration of the proton after it is released? (b) Sketch qualitative (no numbers!) acceleration–time and velocity–time graphs of the released proton's motion.

21.12 • A negative charge of -0.550 μC exerts an upward 0.600 N force on an unknown charge that is located 0.300 m directly below the first charge. What are (a) the value of the unknown charge (magnitude and sign); (b) the magnitude and direction of the force that the unknown charge exerts on the -0.550 μC charge?

21.13 •• Three point charges are arranged on a line. Charge $q_3 = +5.00$ nC and is at the origin. Charge $q_2 = -3.00$ nC and is at $x = +4.00$ cm. Charge q_1 is at $x = +2.00$ cm. What is q_1 (magnitude and sign) if the net force on q_3 is zero?

21.14 •• In Example 21.4, suppose the point charge on the y-axis at $y = -0.30$ m has negative charge -2.0 μC, and the other charges remain the same. Find the magnitude and direction of the net force on Q. How does your answer differ from that in Example 21.4? Explain the differences.

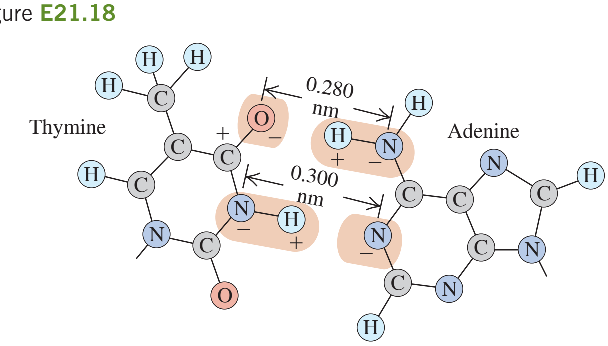
21.15 •• In Example 21.3, calculate the net force on charge q_1.

21.16 •• Three point charges are arranged along the x-axis. Charge $q_1 = +3.00$ μC is at the origin, and charge $q_2 = -5.00$ μC is at $x = 0.200$ m. Charge $q_3 = -8.00$ μC. Where is q_3 located if the net force on q_1 is 7.00 N in the $-x$-direction?

21.17 •• Two point charges are located on the y-axis as follows: charge $q_1 = -1.50$ nC at $y = -0.600$ m, and charge $q_2 = +3.20$ nC at the origin ($y = 0$). What is the total force (magnitude and direction) exerted by these two charges on a third charge $q_3 = +5.00$ nC located at $y = -0.400$ m?

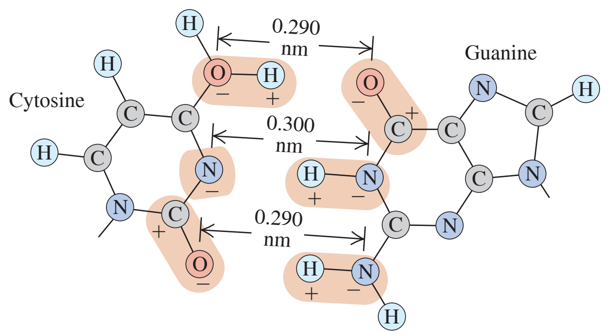
21.18 •• BIO **Base Pairing in DNA, I.** The two sides of the DNA double helix are connected by pairs of bases (adenine, thymine, cytosine, and guanine). Because of the geometric shape of these molecules, adenine bonds with thymine and cytosine bonds with guanine. **Figure E21.18** shows the bonding of thymine and adenine. Each charge shown is $\pm e$, and the H—N distance is 0.110 nm. (a) Calculate the *net* force that thymine exerts on adenine. Is it attractive or repulsive? To keep the calculations fairly simple, yet reasonable, consider only the forces due to the O—H—N and the N—H—N combinations, assuming that these two combinations are parallel to each other. Remember, however, that in the O—H—N set, the O^- exerts a force on both the H^+ and the N^-, and likewise along the N—H—N set. (b) Calculate the force on the electron in the hydrogen atom, which is 0.0529 nm from the proton. Then compare the strength of the bonding force of the electron in hydrogen with the bonding force of the adenine–thymine molecules.

Figure **E21.18**



21.19 •• **BIO** **Base Pairing in DNA, II.** Refer to Exercise 21.18. **Figure E21.19** shows the bonding of cytosine and guanine. The O—H and H—N distances are each 0.110 nm. In this case, assume that the bonding is due only to the forces along the O—H—O, N—H—N, and O—H—N combinations, and assume also that these three combinations are parallel to each other. Calculate the *net* force that cytosine exerts on guanine due to the preceding three combinations. Is this force attractive or repulsive?

Figure **E21.19**



Section 21.4 Electric Field and Electric Forces

21.20 •• **CP** A uniform electric field exists in the region between two oppositely charged plane parallel plates. A proton is released from rest at the surface of the positively charged plate and strikes the surface of the opposite plate, 1.60 cm distant from the first, in a time interval of 3.20×10^{-6} s. (a) Find the magnitude of the electric field. (b) Find the speed of the proton when it strikes the negatively charged plate.

21.21 • **CP** A proton is placed in a uniform electric field of 2.75×10^3 N/C. Calculate (a) the magnitude of the electric force felt by the proton; (b) the proton's acceleration; (c) the proton's speed after 1.00 μs in the field, assuming it starts from rest.

21.22 • A particle has charge -5.00 nC. (a) Find the magnitude and direction of the electric field due to this particle at a point 0.250 m directly above it. (b) At what distance from this particle does its electric field have a magnitude of 12.0 N/C?

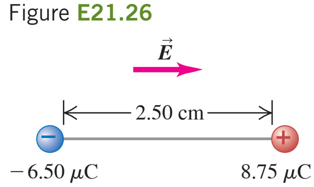
21.23 • **CP** A proton is traveling horizontally to the right at 4.50×10^6 m/s. (a) Find the magnitude and direction of the weakest electric field that can bring the proton uniformly to rest over a distance of 3.20 cm. (b) How much time does it take the proton to stop after entering the field? (c) What minimum field (magnitude and direction) would be needed to stop an electron under the conditions of part (a)?

21.24 • **CP** An electron is released from rest in a uniform electric field. The electron accelerates vertically upward, traveling 4.50 m in the first 3.00 μs after it is released. (a) What are the magnitude and direction of the electric field? (b) Are we justified in ignoring the effects of gravity? Justify your answer quantitatively.

21.25 •• (a) What must the charge (sign and magnitude) of a 1.45 g particle be for it to remain stationary when placed in a downward-directed electric field of magnitude 650 N/C? (b) What is the magnitude of an electric field in which the electric force on a proton is equal in magnitude to its weight?

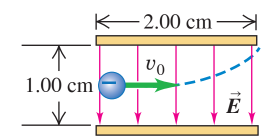
21.26 •• A $+8.75$ μC point charge is glued down on a horizontal frictionless table. It is tied to a -6.50 μC point charge by a light, nonconducting 2.50 cm wire. A uniform electric field of magnitude 1.85×10^8 N/C is directed parallel to the wire, as shown in **Fig. E21.26**. (a) Find the tension in the wire. (b) What would the tension be if both charges were negative?

Figure **E21.26**



21.27 •• **CP** An electron is projected with an initial speed $v_0 = 1.60 \times 10^6$ m/s into the uniform field between two parallel plates (**Fig. E21.27**). Assume that the field between the plates is uniform and directed vertically downward and that the field outside the plates is zero. The electron enters the field at a point midway between the plates. (a) If the electron just misses the upper plate as it emerges from the field, find the magnitude of the electric field. (b) Suppose that the electron in Fig. E21.27 is replaced by a proton with the same initial speed v_0. Would the proton hit one of the plates? If not, what would be the magnitude and direction of its vertical displacement as it exits the region between the plates? (c) Compare the paths traveled by the electron and the proton, and explain the differences. (d) Discuss whether it is reasonable to ignore the effects of gravity for each particle.

Figure **E21.27**



21.28 •• **CP** In Exercise 21.27, what is the speed of the electron as it emerges from the field?

21.29 • A point charge is at the origin. With this point charge as the source point, what is the unit vector $\hat{r}$ in the direction of the field point (a) at $x = 0$, $y = -1.35$ m; (b) at $x = 12.0$ cm, $y = 12.0$ cm; (c) at $x = -1.10$ m, $y = 2.60$ m? Express your results in terms of the unit vectors $\hat{\imath}$ and $\hat{\jmath}$.

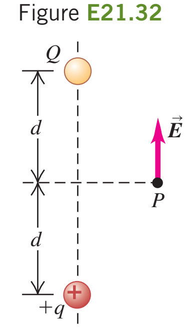
21.30 •• (a) An electron is moving east in a uniform electric field of 1.50 N/C directed to the west. At point *A*, the velocity of the electron is 4.50×10^5 m/s toward the east. What is the speed of the electron when it reaches point *B*, 0.375 m east of point *A*? (b) A proton is moving in the uniform electric field of part (a). At point *A*, the velocity of the proton is 1.90×10^4 m/s, east. What is the speed of the proton at point *B*?

Section 21.5 Electric-Field Calculations

21.31 •• A uniform line of charge with length 20.0 cm is along the *x*-axis, with its midpoint at $x = 0$. Its charge per length is $+4.80$ nC/m. A small sphere with charge -2.00 μC is located at $x = 0$, $y = 5.00$ cm. What are the magnitude and direction of the force that the charged sphere exerts on the line of charge?

21.32 • Two point charges Q and $+q$ (where q is positive) produce the net electric field shown at point *P* in **Fig. E21.32**. The field points parallel to the line connecting the two charges. (a) What can you conclude about the sign and magnitude of Q? Explain your reasoning. (b) If the lower charge were negative instead, would it be possible for the field to have the direction shown in the figure? Explain your reasoning.

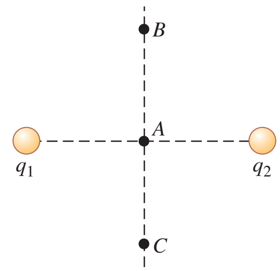
Figure **E21.32**



21.33 •• A very long line of charge with charge per unit length $+8.00$ μC/m is on the *x*-axis and its midpoint is at $x = 0$. A second very long line of charge with charge per length -4.00 μC/m is parallel to the *x*-axis at $y = 10.0$ cm and its midpoint is also at $x = 0$. At what point on the *y*-axis is the resultant electric field of the two lines of charge equal to zero?

21.34 • The two charges q_1 and q_2 shown in **Fig. E21.34** have equal magnitudes. What is the direction of the net electric field due to these two charges at points *A* (midway between the charges), *B*, and *C* if (a) both charges are negative, (b) both charges are positive, (c) q_1 is positive and q_2 is negative.

Figure **E21.34**

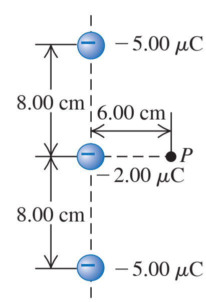


21.35 • A +2.00 nC point charge is at the origin, and a second −5.00 nC point charge is on the x-axis at $x = 0.800$ m. (a) Find the electric field (magnitude and direction) at each of the following points on the x-axis: (i) $x = 0.200$ m; (ii) $x = 1.20$ m; (iii) $x = -0.200$ m. (b) Find the net electric force that the two charges would exert on an electron placed at each point in part (a).

21.36 •• Repeat Exercise 21.35, but now let the charge at the origin be −4.00 nC.

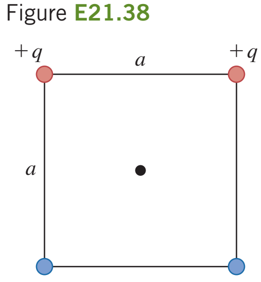
21.37 • Three negative point charges lie along a line as shown in **Fig. E21.37**. Find the magnitude and direction of the electric field this combination of charges produces at point P, which lies 6.00 cm from the −2.00 μC charge measured perpendicular to the line connecting the three charges.

Figure **E21.37**



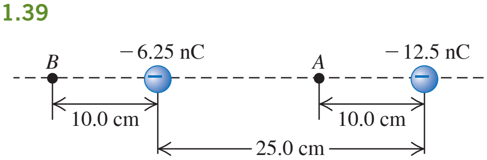
21.38 •• A point charge is placed at each corner of a square with side length a. All charges have magnitude q. Two of the charges are positive and two are negative (**Fig. E21.38**). What is the direction of the net electric field at the center of the square due to the four charges, and what is its magnitude in terms of q and a?

Figure **E21.38**



21.39 • Two point charges are separated by 25.0 cm (**Fig. E21.39**). Find the net electric field these charges produce at (a) point A and (b) point B. (c) What would be the magnitude and direction of the electric force this combination of charges would produce on a proton at A?

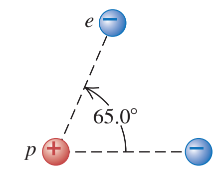
Figure **E21.39**



21.40 •• Point charge $q_1 = -5.00$ nC is at the origin and point charge $q_2 = +3.00$ nC is on the x-axis at $x = 3.00$ cm. Point P is on the y-axis at $y = 4.00$ cm. (a) Calculate the electric fields $\vec{E}_1$ and $\vec{E}_2$ at point P due to the charges q_1 and q_2. Express your results in terms of unit vectors (see Example 21.6). (b) Use the results of part (a) to obtain the resultant field at P, expressed in unit vector form.

21.41 •• If two electrons are each 1.50×10^{-10} m from a proton (**Fig. E21.41**), find the magnitude and direction of the net electric force they will exert on the proton.

Figure **E21.41**



21.42 •• **BIO Electric Field of Axons.** A nerve signal is transmitted through a neuron when an excess of Na^+ ions suddenly enters the axon, a long cylindrical part of the neuron. Axons are approximately 10.0 μm in diameter, and measurements show that about 5.6×10^{11} Na^+ ions per meter (each of charge $+e$) enter during this process. Although the axon is a long cylinder, the charge does not all enter everywhere at the same time. A plausible model would be a series of point charges moving along the axon. Consider a 0.10 mm length of the axon and model it as a point charge. (a) If the charge that enters each meter of the axon gets distributed uniformly along it, how many coulombs of charge enter a 0.10 mm length of the axon? (b) What electric field (magnitude and direction) does the sudden influx of charge produce at the surface of the body if the axon is 5.00 cm below the skin? (c) Certain sharks can respond to electric fields as weak as 1.0 μN/C. How far from this segment of axon could a shark be and still detect its electric field?

21.43 • In a rectangular coordinate system a positive point charge $q = 6.00 \times 10^{-9}$ C is placed at the point $x = +0.150$ m, $y = 0$, and an identical point charge is placed at $x = -0.150$ m, $y = 0$. Find the x- and y-components, the magnitude, and the direction of the electric field at the following points: (a) the origin; (b) $x = 0.300$ m, $y = 0$; (c) $x = 0.150$ m, $y = -0.400$ m; (d) $x = 0$, $y = 0.200$ m.

21.44 •• A point charge $q_1 = -4.00$ nC is at the point $x = 0.600$ m, $y = 0.800$ m, and a second point charge $q_2 = +6.00$ nC is at the point $x = 0.600$ m, $y = 0$. Calculate the magnitude and direction of the net electric field at the origin due to these two point charges.

21.45 • Three parallel sheets of charge, large enough to be treated as infinite sheets, are perpendicular to the x-axis. Sheet A has surface charge density $\sigma_A = +8.00\ \text{nC/m}^2$. Sheet B is 4.00 cm to the right of sheet A and has surface charge density $\sigma_B = -4.00\ \text{nC/m}^2$. Sheet C is 4.00 cm to the right of sheet B, so is 8.00 cm to the right of sheet A, and has surface charge density $\sigma_C = +6.00\ \text{nC/m}^2$. What are the magnitude and direction of the resultant electric field at a point that is midway between sheets B and C, or 2.00 cm from each of these two sheets?

21.46 • Point charge $q_1 = -5.00$ nC is on the x-axis at $x = -0.400$ m. Point P is on the x-axis at $x = +0.200$ m. Point charge q_2 is at the origin. What are the sign and magnitude of q_2 if the resultant electric field at point P is zero?

21.47 • Point charge A is on the x-axis at $x = -3.00$ cm. At $x = 1.00$ cm on the x-axis its electric field is 2700 N/C. Point charge B is also on the x-axis, at $x = 5.00$ cm. The absolute magnitude of charge B is twice that of A. Find the magnitude and direction of the total electric field at the origin if (a) both A and B are positive; (b) both are negative; (c) A is positive and B is negative; (d) A is negative and B is positive.

21.48 •• A very long, straight wire has charge per unit length 3.20×10^{-10} C/m. At what distance from the wire is the electric-field magnitude equal to 2.50 N/C?

21.49 • A ring-shaped conductor with radius $a = 2.50$ cm has a total positive charge $Q = +0.125$ nC uniformly distributed around it (see Fig. 21.23). The center of the ring is at the origin of coordinates O. (a) What is the electric field (magnitude and direction) at point P, which is on the x-axis at $x = 40.0$ cm? (b) A point charge $q = -2.50$ μC is placed at P. What are the magnitude and direction of the force exerted *by* the charge q *on* the ring?

21.50 •• A straight, nonconducting plastic wire 8.50 cm long carries a charge density of +175 nC/m distributed uniformly along its length. It is lying on a horizontal tabletop. (a) Find the magnitude and direction of the electric field this wire produces at a point 6.00 cm directly above its midpoint. (b) If the wire is now bent into a circle lying flat on the table, find the magnitude and direction of the electric field it produces at a point 6.00 cm directly above its center.

Section 21.7 Electric Dipoles

21.51 • Point charges $q_1 = -4.5$ nC and $q_2 = +4.5$ nC are separated by 3.1 mm, forming an electric dipole. (a) Find the electric dipole moment (magnitude and direction). (b) The charges are in a uniform electric field whose direction makes an angle of 36.9° with the line connecting the charges. What is the magnitude of this field if the torque exerted on the dipole has magnitude 7.2×10^{-9} N·m?

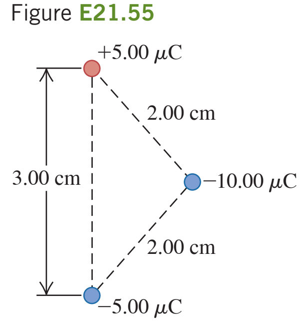
21.52 • The ammonia molecule (NH_3) has a dipole moment of 5.0×10^{-30} C·m. Ammonia molecules in the gas phase are placed in a uniform electric field $\vec{E}$ with magnitude 1.6×10^6 N/C. (a) What is the change in electric potential energy when the dipole moment of a molecule changes its orientation with respect to $\vec{E}$ from parallel to perpendicular? (b) At what absolute temperature T is the average translational kinetic energy $\frac{3}{2}kT$ of a molecule equal to the change in potential energy calculated in part (a)? (*Note:* Above this temperature, thermal agitation prevents the dipoles from aligning with the electric field.)

21.53 • **Torque on a Dipole.** An electric dipole with dipole moment $\vec{p}$ is in a uniform external electric field $\vec{E}$. (a) Find the orientations of the dipole for which the torque on the dipole is zero. (b) Which of the orientations in part (a) is stable, and which is unstable? (*Hint:* Consider a small rotation away from the equilibrium position and see what happens.) (c) Show that for the stable orientation in part (b), the dipole's own electric field tends to oppose the external field.

21.54 • The dipole moment of the water molecule (H_2O) is 6.17×10^{-30} C·m. Consider a water molecule located at the origin whose dipole moment $\vec{p}$ points in the +x-direction. A chlorine ion ($C1^-$), of charge -1.60×10^{-19} C, is located at $x = 3.00 \times 10^{-9}$ m. Find the magnitude and direction of the electric force that the water molecule exerts on the chlorine ion. Is this force attractive or repulsive? Assume that x is much larger than the separation d between the charges in the dipole, so that the approximate expression for the electric field along the dipole axis derived in Example 21.14 can be used.

21.55 • Three charges are at the corners of an isosceles triangle as shown in **Fig. E21.55**. The $\pm 5.00\ \mu$C charges form a dipole. (a) Find the force (magnitude and direction) the $-10.00\ \mu$C charge exerts on the dipole. (b) For an axis perpendicular to the line connecting the $\pm 5.00\ \mu$C charges at the midpoint of this line, find the torque (magnitude and direction) exerted on the dipole by the $-10.00\ \mu$C charge.

Figure E21.55



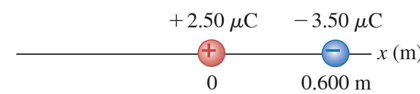
21.56 • An electric dipole with a dipole moment of magnitude p is placed at various orientations in an electric field $\vec{E}$ that is directed to the left. (a) What orientation of the dipole will result in maximum torque directed into the page? What then is the electric potential energy? (b) What orientation of the dipole will give zero torque and maximum electric potential energy? What type of equilibrium is this: stable, unstable, or neutral?

PROBLEMS

21.57 ••• Four identical charges Q are placed at the corners of a square of side L. (a) In a free-body diagram, show all of the forces that act on one of the charges. (b) Find the magnitude and direction of the total force exerted on one charge by the other three charges.

21.58 ••• Two charges are placed on the x-axis: one, of 2.50 μC, at the origin and the other, of $-3.50\ \mu$C, at $x = 0.600$ m (**Fig. P21.58**). Find the position on the x-axis where the net force on a small charge $+q$ would be zero.

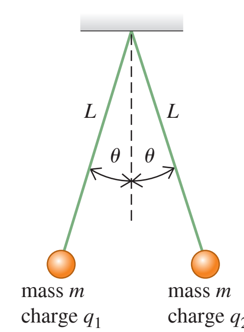
Figure P21.58



21.59 •• A charge $q_1 = +5.00$ nC is placed at the origin of an xy-coordinate system, and a charge $q_2 = -2.00$ nC is placed on the positive x-axis at $x = 4.00$ cm. (a) If a third charge $q_3 = +6.00$ nC is now placed at the point $x = 4.00$ cm, $y = 3.00$ cm, find the x- and y-components of the total force exerted on this charge by the other two. (b) Find the magnitude and direction of this force.

21.60 •• **CP** Two identical spheres with mass m are hung from silk threads of length L (**Fig. P21.60**). The spheres have the same charge, so $q_1 = q_2 = q$. The radius of each sphere is very small compared to the distance between the spheres, so they may be treated as point charges. Show that if the angle θ is small, the equilibrium separation d between the spheres is $d = (q^2L/2\pi\epsilon_0 mg)^{1/3}$. (*Hint:* If θ is small, then $\tan\theta \cong \sin\theta$.)

Figure P21.60

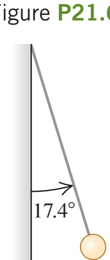


21.61 ••• **CP** Two small spheres with mass $m = 15.0$ g are hung by silk threads of length $L = 1.20$ m from a common point (Fig. P21.60). When the spheres are given equal quantities of negative charge, so that $q_1 = q_2 = q$, each thread hangs at $\theta = 25.0°$ from the vertical. (a) Draw a diagram showing the forces on each sphere. Treat the spheres as point charges. (b) Find the magnitude of q. (c) Both threads are now shortened to length $L = 0.600$ m, while the charges q_1 and q_2 remain unchanged. What new angle will each thread make with the vertical? (*Hint:* This part of the problem can be solved numerically by using trial values for θ and adjusting the values of θ until a self-consistent answer is obtained.)

21.62 •• **CP** A small sphere with charge $q = 5.00\ \mu$C and mass 0.500 g is traveling horizontally toward the east at a height of 60.0 cm above the ground. The sphere has a speed of 2.00 m/s as it enters a region of uniform electric field with magnitude E. What is E if the sphere has a speed of 5.00 m/s just before it strikes the ground?

21.63 •• **CP** A small 12.3 g plastic ball is tied to a very light 28.6 cm string that is attached to the vertical wall of a room (**Fig. P21.63**). A uniform horizontal electric field exists in this room. When the ball has been given an excess charge of $-1.11\ \mu$C, you observe that it remains suspended, with the string making an angle of 17.4° with the wall. Find the magnitude and direction of the electric field in the room.

Figure P21.63



21.64 •• **CP** A small sphere with positive charge q and mass m is released from rest in a uniform electric field $\vec{E}$ that is directed vertically upward. The magnitude of the field is large enough for the sphere to travel upward when it is released. How long does it take the sphere to travel upward a distance d after being released from rest? Give your answer in terms of q, m, d, E, and the acceleration due to gravity, g.

21.65 •• If we rub a balloon on our hair, the balloon sticks to a wall or ceiling. This is because the rubbing transfers electrons from our hair to the balloon, giving it a net negative charge. When the balloon is placed near the ceiling, the extra electrons in it repel nearby electrons in the ceiling, creating a separation of charge on the ceiling, with positive charge closer to the balloon. Model the interaction as two point-like charges of equal magnitude and opposite signs, separated by a distance of 500 μm. Neglect the more distant negative charges on the ceiling. (a) A typical balloon has a mass of 4 g. Estimate the minimum magnitude of charge the balloon requires to stay attached to the ceiling. (b) Since a balloon sticks handily to the ceiling after being rubbed, assume that it has attained 10 times the estimated minimum charge. Estimate the number of electrons that were transferred to the balloon by the process of rubbing.

21.66 •• An American penny is 97.5% zinc and 2.5% copper and has a mass of 2.5 g. (a) Use the approximation that a penny is pure zinc, which has an atomic mass of 65.38 g/mol, to estimate the number of electrons in a penny. (Each zinc atom has 30 electrons.) (b) Estimate the net charge on all of the electrons in one penny. (c) The net positive charge on all of the protons in a penny has the same magnitude as the charge on the electrons. Estimate the force on either of two objects with this net magnitude of charge if the objects are separated by 2 cm. (d) Estimate the number of leaves on an oak tree that is 60 feet tall. (e) Imagine a forest filled with such trees, arranged in a square lattice, each 10 m distant from its neighbors. Estimate how large such a forest would need to be to include as many leaves as there are electrons in one penny. (f) How does that area compare to the surface area of the earth?

21.67 •• Two particles having charges $q_1 = 0.500$ nC and $q_2 = 8.00$ nC are separated by a distance of 1.20 m. At what point along the line connecting the two charges is the total electric field due to the two charges equal to zero?

21.68 •• A -3.00 nC point charge is on the x-axis at $x = 1.20$ m. A second point charge, Q, is on the x-axis at -0.600 m. What must be the sign and magnitude of Q for the resultant electric field at the origin to be (a) 45.0 N/C in the $+x$-direction, (b) 45.0 N/C in the $-x$-direction?

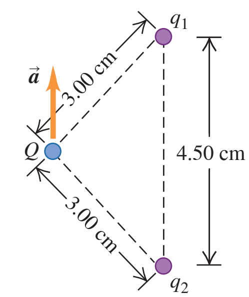
21.69 ••• A charge $+Q$ is located at the origin, and a charge $+4Q$ is at distance d away on the x-axis. Where should a third charge, q, be placed, and what should be its sign and magnitude, so that all three charges will be in equilibrium?

21.70 •• A charge of -3.00 nC is placed at the origin of an xy-coordinate system, and a charge of 2.00 nC is placed on the y-axis at $y = 4.00$ cm. (a) If a third charge, of 5.00 nC, is now placed at the point $x = 3.00$ cm, $y = 4.00$ cm, find the x- and y-components of the total force exerted on this charge by the other two charges. (b) Find the magnitude and direction of this force.

21.71 • Three identical point charges q are placed at each of three corners of a square of side L. Find the magnitude and direction of the net force on a point charge $-3q$ placed (a) at the center of the square and (b) at the vacant corner of the square. In each case, draw a free-body diagram showing the forces exerted on the $-3q$ charge by each of the other three charges.

21.72 ••• Two point charges q_1 and q_2 are held in place 4.50 cm apart. Another point charge $Q = -1.75$ μC, of mass 5.00 g, is initially located 3.00 cm from both of these charges (**Fig. P21.72**) and released from rest. You observe that the initial acceleration of Q is 324 m/s^2 upward, parallel to the line connecting the two point charges. Find q_1 and q_2.

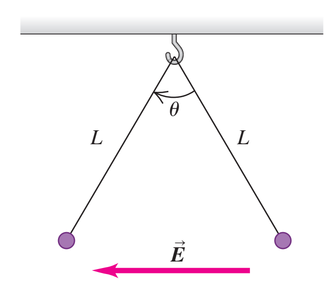
Figure **P21.72**



21.73 •• **CP** **Strength of the Electric Force.** Imagine two 1.0 g bags of protons, one at the earth's north pole and the other at the south pole. (a) How many protons are in each bag? (b) Calculate the gravitational attraction and the electric repulsion that each bag exerts on the other. (c) Are the forces in part (b) large enough for you to feel if you were holding one of the bags?

21.74 ••• **CP** Two tiny spheres of mass 6.80 mg carry charges of equal magnitude, 72.0 nC, but opposite sign. They are tied to the same ceiling hook by light strings of length 0.530 m. When a horizontal uniform electric field E that is directed to the left is turned on, the spheres hang at rest with the angle θ between the strings equal to 58.0° (**Fig. P21.74**). (a) Which ball (the one on the right or the one on the left) has positive charge? (b) What is the magnitude E of the field?

Figure **P21.74**



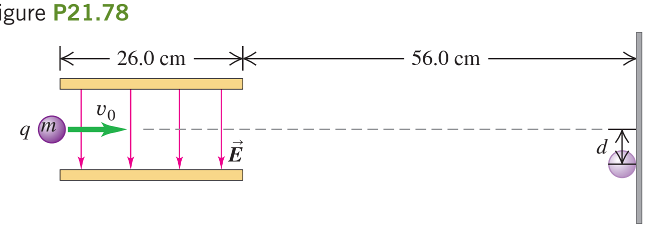
21.75 •• **CP** Consider a model of a hydrogen atom in which an electron is in a circular orbit of radius $r = 5.29 \times 10^{-11}$ m around a stationary proton. What is the speed of the electron in its orbit?

21.76 ••• **CP** A disk with radius R and uniform positive charge density σ lies horizontally on a tabletop. A small plastic sphere with mass M and positive charge Q hovers motionless above the center of the disk, suspended by the Coulomb repulsion due to the charged disk. (a) What is the magnitude of the net upward force on the sphere as a function of the height z above the disk? (b) At what height h does the sphere hover? Express your answer in terms of the dimensionless constant $v \equiv 2\epsilon_0 Mg/(Q\sigma)$. (c) If $M = 100$ g, $Q = 1$ μC, $R = 5$ cm, and $\sigma = 10$ nC/cm^2, what is h?

21.77 •• **CP** A proton is projected into a uniform electric field that points vertically upward and has magnitude E. The initial velocity of the proton has a magnitude v_0 and is directed at an angle α below the horizontal. (a) Find the maximum distance h_{max} that the proton descends vertically below its initial elevation. Ignore gravitational forces. (b) After what horizontal distance d does the proton return to its original elevation? (c) Sketch the trajectory of the proton. (d) Find the numerical values of h_{max} and d if $E = 500$ N/C, $v_0 = 4.00 \times 10^5$ m/s, and $\alpha = 30.0°$.

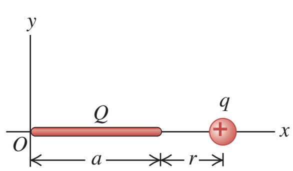
21.78 ••• A small object with mass m, charge q, and initial speed $v_0 = 5.00 \times 10^3$ m/s is projected into a uniform electric field between two parallel metal plates of length 26.0 cm (**Fig. P21.78**). The electric field between the plates is directed downward and has magnitude $E = 800$ N/C. Assume that the field is zero outside the region between the plates. The separation between the plates is large enough for the object to pass between the plates without hitting the lower plate. After passing through the field region, the object is deflected downward a vertical distance $d = 1.25$ cm from its original direction of motion and reaches a collecting plate that is 56.0 cm from the edge of the parallel plates. Ignore gravity and air resistance. Calculate the object's charge-to-mass ratio, q/m.

Figure **P21.78**



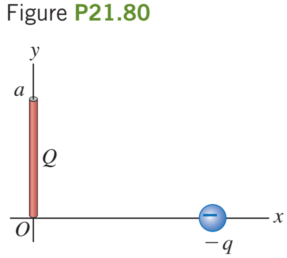
21.79 •• **CALC** Positive charge Q is distributed uniformly along the x-axis from $x = 0$ to $x = a$. A positive point charge q is located on the positive x-axis at $x = a + r$, a distance r to the right of the end of Q (**Fig. P21.79**). (a) Calculate the x- and y-components of the electric field produced by the charge distribution Q at points on the positive x-axis where $x > a$. (b) Calculate the force (magnitude and direction) that the charge distribution Q exerts on q. (c) Show that if $r \gg a$, the magnitude of the force in part (b) is approximately $Qq/4\pi\epsilon_0 r^2$. Explain why this result is obtained.

Figure **P21.79**



21.80 •• **CALC** Positive charge Q is distributed uniformly along the positive y-axis between $y = 0$ and $y = a$. A negative point charge $-q$ lies on the positive x-axis, a distance x from the origin (**Fig. P21.80**). (a) Calculate the x- and y-components of the electric field produced by the charge distribution Q at points on the positive x-axis. (b) Calculate the x- and y-components of the force that the charge distribution Q exerts on q. (c) Show that if $x \gg a$, $F_x \cong -Qq/4\pi\epsilon_0 x^2$ and $F_y \cong +Qqa/8\pi\epsilon_0 x^3$. Explain why this result is obtained.

Figure **P21.80**



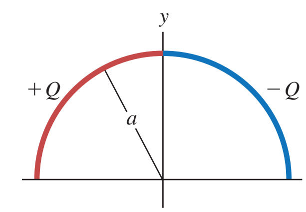
21.81 ••• A uniformly charged disk like the disk in Fig. 21.25 has radius 2.50 cm and carries a total charge of 7.0×10^{-12} C. (a) Find the electric field (magnitude and direction) on the x-axis at $x = 20.0$ cm. (b) Show that for $x \gg R$, Eq. (21.11) becomes $E = Q/4\pi\epsilon_0 x^2$, where Q is the total charge on the disk. (c) Is the magnitude of the electric field you calculated in part (a) larger or smaller than the electric field 20.0 cm from a point charge that has the same total charge as this disk? In terms of the approximation used in part (b) to derive $E = Q/4\pi\epsilon_0 x^2$ for a point charge from Eq. (21.11), explain why this is so. (d) What is the percent difference between the electric fields produced by the finite disk and by a point charge with the same charge at $x = 20.0$ cm and at $x = 10.0$ cm?

21.82 •• **CP** A small sphere with mass m carries a positive charge q and is attached to one end of a silk fiber of length L. The other end of the fiber is attached to a large vertical insulating sheet that has a positive surface charge density σ. Show that when the sphere is in equilibrium, the fiber makes an angle equal to arctan $(q\sigma/2mg\epsilon_0)$ with the vertical sheet.

21.83 •• **CALC** Negative charge $-Q$ is distributed uniformly around a quarter-circle of radius a that lies in the first quadrant, with the center of curvature at the origin. Find the x- and y-components of the net electric field at the origin.

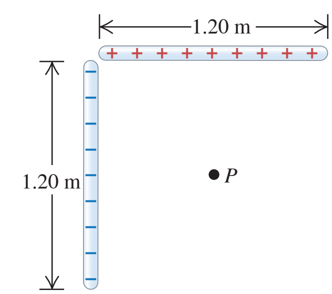
21.84 •• **CALC** A semicircle of radius a is in the first and second quadrants, with the center of curvature at the origin. Positive charge $+Q$ is distributed uniformly around the left half of the semicircle, and negative charge $-Q$ is distributed uniformly around the right half of the semicircle (**Fig. P21.84**). What are the magnitude and direction of the net electric field at the origin produced by this distribution of charge?

Figure **P21.84**



21.85 •• Two 1.20 m nonconducting rods meet at a right angle. One rod carries $+2.50\ \mu$C of charge distributed uniformly along its length, and the other carries $-2.50\ \mu$C distributed uniformly along it (**Fig. P21.85**). (a) Find the magnitude and direction of the electric field these rods produce at point P, which is 60.0 cm from each rod. (b) If an electron is released at P, what are the magnitude and direction of the net force that these rods exert on it?

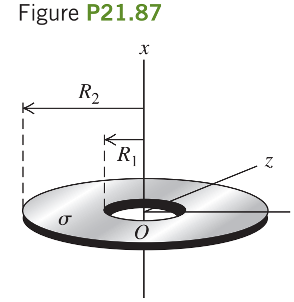
Figure **P21.85**



21.86 • Two very large parallel sheets are 5.00 cm apart. Sheet A carries a uniform surface charge density of $-8.80\ \mu\text{C/m}^2$, and sheet B, which is to the right of A, carries a uniform charge density of $-11.6\ \mu\text{C/m}^2$. Assume that the sheets are large enough to be treated as infinite. Find the magnitude and direction of the net electric field these sheets produce at a point (a) 4.00 cm to the right of sheet A; (b) 4.00 cm to the left of sheet A; (c) 4.00 cm to the right of sheet B.

21.87 •• **CP** A thin disk with a circular hole at its center, called an *annulus*, has inner radius R_1 and outer radius R_2 (**Fig. P21.87**). The disk has a uniform positive surface charge density σ on its surface. (a) Determine the total electric charge on the annulus. (b) The annulus lies in the yz-plane, with its center at the origin. For an arbitrary point on the x-axis (the axis of the annulus), find the magnitude and direction of the electric field $\vec{E}$. Consider points both above and below the annulus. (c) Show that at points on the x-axis that are sufficiently close to the origin, the magnitude of the electric field is approximately proportional to the distance between the center of the annulus and the point. How close is "sufficiently close"? (d) A point particle with mass m and negative charge $-q$ is free to move along the x-axis (but cannot move off the axis). The particle is originally placed at rest at $x = 0.01 R_1$ and released. Find the frequency of oscillation of the particle. (*Hint:* Review Section 14.2. The annulus is held stationary.)

Figure **P21.87**

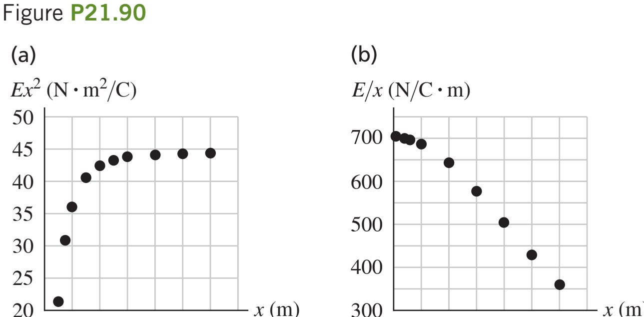


21.88 ••• DATA CP Design of an Inkjet Printer. Inkjet printers can be described as either continuous or drop-on-demand. In a continuous inkjet printer, letters are built up by squirting drops of ink at the paper from a rapidly moving nozzle. You are part of an engineering group working on the design of such a printer. Each ink drop will have a mass of 1.4×10^{-8} g. The drops will leave the nozzle and travel toward the paper at 50 m/s in a horizontal direction, passing through a charging unit that gives each drop a positive charge q by removing some electrons from it. The drops will then pass between parallel deflecting plates, 2.0 cm long, where there is a uniform vertical electric field with magnitude 8.0×10^4 N/C. Your team is working on the design of the charging unit that places the charge on the drops. (a) If a drop is to be deflected 0.30 mm by the time it reaches the end of the deflection plates, what magnitude of charge must be given to the drop? How many electrons must be removed from the drop to give it this charge? (b) If the unit that produces the stream of drops is redesigned so that it produces drops with a speed of 25 m/s, what q value is needed to achieve the same 0.30 mm deflection?

21.89 •• DATA Two small spheres, each carrying a net positive charge, are separated by 0.400 m. You have been asked to perform measurements that will allow you to determine the charge on each sphere. You set up a coordinate system with one sphere (charge q_1) at the origin and the other sphere (charge q_2) at $x = +0.400$ m. Available to you are a third sphere with net charge $q_3 = 4.00 \times 10^{-6}$ C and an apparatus that can accurately measure the location of this sphere and the net force on it. First you place the third sphere on the x-axis at $x = 0.200$ m; you measure the net force on it to be 4.50 N in the $+x$-direction. Then you move the third sphere to $x = +0.600$ m and measure the net force on it now to be 3.50 N in the $+x$-direction. (a) Calculate q_1 and q_2. (b) What is the net force (magnitude and direction) on q_3 if it is placed on the x-axis at $x = -0.200$ m? (c) At what value of x (other than $x = \pm\infty$) could q_3 be placed so that the net force on it is zero?

21.90 ••• DATA Positive charge Q is distributed uniformly around a very thin conducting ring of radius a, as in Fig. 21.23. You measure the electric field E at points on the ring axis, at a distance x from the center of the ring, over a wide range of values of x. (a) Your results for the larger values of x are plotted in **Fig. P21.90a** as Ex^2 versus x. Explain why the quantity Ex^2 approaches a constant value as x increases. Use Fig. P21.90a to calculate the net charge Q on the ring. (b) Your results for smaller values of x are plotted in Fig. P21.90b as E/x versus x. Explain why E/x approaches a constant value as x approaches zero. Use Fig. P21.90b to calculate a.

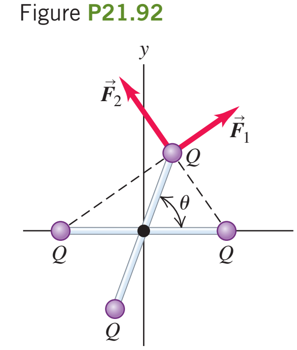
Figure **P21.90**



21.91 ••• CP Consider an infinite flat sheet with positive charge density σ in which a circular hole of radius R has been cut out. The sheet lies in the xy-plane with the origin at the center of the hole. The sheet is parallel to the ground, so that the positive z-axis describes the "upward" direction. If a particle of mass m and negative charge $-q$ sits at rest at the center of the hole and is released, the particle, constrained to the z-axis, begins to fall. As it drops farther beneath the sheet, the upward electric force increases. For a sufficiently low value of m, the upward electrical attraction eventually exceeds the particle's weight and the particle will slow, come to a stop, and then rise back to its original position. This sequence of events will repeat indefinitely. (a) What is the electric field at a depth Δ beneath the origin along the negative z-axis? (b) What is the maximum mass m_{max} that would prevent the particle from falling indefinitely? (c) If $m < m_{max}$, how much work is done by the electric field as the particle drops from $z = 0$ to $z = -\Delta$? (d) In this same interval, how much work is done by gravity? (e) To what ultimate depth Δ_{max} will the particle drop? Express your answer in terms of the dimensionless parameter $\alpha \equiv m/m_{max}$. (Use the work–energy theorem to find the speed as a function of depth, and then solve for the depth at which the speed is zero.) (f) What is the particle's speed at depth $\Delta \le \Delta_{max}$? (g) If the sheet has charge density 1.00 nC/cm^2, the radius of the hole is $R = 10.0$ cm, and the particle has mass 25.0 g and charge 1.00 μC, what are m_{max} and Δ_{max}?

21.92 ••• CP Two rigid insulating rods, each of length L and negligible mass, are connected at their centers by a frictionless hinge. One rod is fixed horizontally atop a vertical shaft, while the other rod is free to pivot about the hinge horizontally. A small ball with charge Q and mass M is affixed to each end of each rod. When the rods are perpendicular, there is no net torque on the second rod. We may describe the configuration of this system by the angle θ between the two rods. The first rod lies on the x-axis and the hinge sits at the origin. The second rod is rotated clockwise as seen from above in **Fig. 21.92**. (a) Determine the force $\vec{F}_1$ exerted by the charge on the left side of the fixed rod on the charge at the upper right side of the movable rod, as a function of θ. (b) Determine the force $\vec{F}_2$ exerted by the charge on the right side of the fixed rod on the charge at the upper right side of the movable rod, as a function of θ. (c) Determine the torque about the pivot point associated with the force $\vec{F}_1$. (d) Determine the torque about the pivot point associated with the force $\vec{F}_2$. (e) Given that the net torque exerted on the movable rod has contributions from the charges at both ends, write a formula for the net torque on the movable rod. (f) The equilibrium configuration at which the torque vanishes is $\theta = \pi/2$. Deviations from equilibrium may be parameterized as $\theta = \pi/2 - \epsilon$. Using power series expansions (see Appendix B), derive the torque for small ϵ, keeping only the lowest order term. (g) For small amplitudes, this system describes a torsional oscillator. Using the previous result, write an expression for the frequency of the small oscillations.

Figure **P21.92**



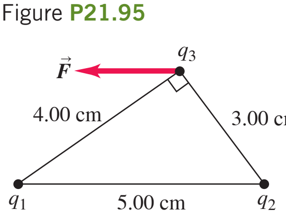
CHALLENGE PROBLEMS

21.93 ••• CP CALC Two thin rods, each with length L and total charge $+Q$, are parallel and separated by a distance a. The first rod has one end at the origin and its other end on the positive y-axis. The second rod has its lower end on the positive x-axis. (a) Explain why the y-component of the net force on the second rod vanishes. (b) Determine the x-component of the differential force dF_2 exerted on a small portion of the second rod, with length dy_2 and position y_2, by the first rod. (This requires integrating over differential portions of the first rod, parameterized by dy_1.) (c) Determine the net force $\vec{F}_2$ on the second rod by integrating dF_{2x} over the second rod. (d) Show that in the limit $a \gg L$ the force determined in part (c) becomes $\frac{1}{4\pi\epsilon_0}\frac{Q^2}{a^2}\hat{\imath}$. (e) Determine the external work required to move the second rod from very far away to the position $x = a$, provided the first rod is held fixed at $x = 0$. This describes the potential energy of the original configuration. (f) Suppose $L = 50.0$ cm, $a = 10.0$ cm, $Q = 10.0\ \mu$C, and $m = 500$ g. If the two rods are released from the original configuration, they will fly apart and ultimately achieve a particular *relative speed*. What is that relative speed?

21.94 ••• CP CALC An insulating rigid rod of length $2a$ and negligible mass is attached at its center to a pivot at the origin and is free to rotate in the xy-plane. A small ball with mass M and charge Q is attached to one end of the rod. A second small ball with mass M and no charge is attached to the other end. A constant electric field $\vec{E} = -E\hat{\imath}$ is present in the region $y > 0$, while the region $y < 0$ has a vanishing electric field. Define $\vec{r}$ as the vector that points from the center of the rod to the charged end of the rod, and θ as the angle between $\vec{r}$ and the positive x-axis. The rod is oriented so that $\theta = 0$ and is given an infinitesimal nudge in the direction of increasing θ. (a) Write an expression for the vector $\vec{r}$. (b) Determine the torque $\vec{\tau}$ about the center of the rod when $0 \le \theta \le \pi$. (c) Determine the torque on the rod about its center when $\pi \le \theta \le 2\pi$. (d) What is the moment of inertia I of the system about the z-axis? (e) The potential energy $U(\theta)$ is determined by $\tau = -dU/d\theta$. Use this equation to write an expression for $U(\theta)$ over the range $0 \le \theta \le 4\pi$, using the convention that $U(0) = 0$. Make sure that $U(\theta)$ is continuous. (f) The angular velocity of the rod is $\omega = \omega(\theta)$. Using $\tau = I\,d^2\theta/dt^2$, show that the energy $\frac{1}{2}I\omega^2 + U(\theta)$ is conserved. (g) Using energy conservation, determine an expression for the angular velocity at the nth time the positive charge crosses the negative y-axis.

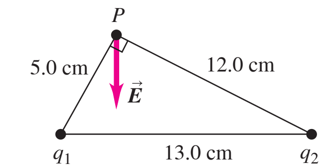
21.95 ••• Three charges are placed as shown in **Fig. P21.95**. The magnitude of q_1 is 2.00 μC, but its sign and the value of the charge q_2 are not known. Charge q_3 is $+4.00\ \mu$C, and the net force $\vec{F}$ on q_3 is entirely in the negative x-direction. (a) Considering the different possible signs of q_1, there are four possible force diagrams representing the forces $\vec{F}_1$ and $\vec{F}_2$ that q_1 and q_2 exert on q_3. Sketch these four possible force configurations. (b) Using the sketches from part (a) and the direction of $\vec{F}$, deduce the signs of the charges q_1 and q_2. (c) Calculate the magnitude of q_2. (d) Determine F, the magnitude of the net force on q_3.

Figure P21.95



21.96 ••• Two charges are placed as shown in **Fig. P21.96**. The magnitude of q_1 is 3.00 μC, but its sign and the value of the charge q_2 are not known. The direction of the net electric field $\vec{E}$ at point P is entirely in the negative y-direction. (a) Considering the different possible signs of q_1 and q_2, four possible diagrams could represent the electric fields $\vec{E}_1$ and $\vec{E}_2$ produced by q_1 and q_2. Sketch the four possible electric-field configurations. (b) Using the sketches from part (a) and the direction of $\vec{E}$, deduce the signs of q_1 and q_2. (c) Determine the magnitude of $\vec{E}$.

Figure P21.96



21.97 ••• CALC Two thin rods of length L lie along the x-axis, one between $x = \frac{1}{2}a$ and $x = \frac{1}{2}a + L$ and the other between $x = -\frac{1}{2}a$ and $x = -\frac{1}{2}a - L$. Each rod has positive charge Q distributed uniformly along its length. (a) Calculate the electric field produced by the second rod at points along the positive x-axis. (b) Show that the magnitude of the force that one rod exerts on the other is

$$F = \frac{Q^2}{4\pi\epsilon_0 L^2}\ln\left[\frac{(a + L)^2}{a(a + 2L)}\right]$$

(c) Show that if $a \gg L$, the magnitude of this force reduces to $F = Q^2/4\pi\epsilon_0 a^2$. (*Hint:* Use the expansion $\ln(1 + z) = z - \frac{1}{2}z^2 + \frac{1}{3}z^3 - \cdots$, valid for $|z| \ll 1$. Carry *all* expansions to at least order L^2/a^2.) Interpret this result.

MCAT-STYLE PASSAGE PROBLEMS

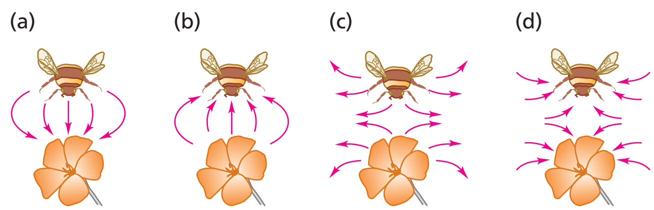
BIO **Electric Bees.** Flying insects such as bees may accumulate a small positive electric charge as they fly. In one experiment, the mean electric charge of 50 bees was measured to be $+(30 \pm 5)$ pC per bee. Researchers also observed the electrical properties of a plant consisting of a flower atop a long stem. The charge on the stem was measured as a positively charged bee approached, landed, and flew away. Plants are normally electrically neutral, so the measured net electric charge on the stem was zero when the bee was very far away. As the bee approached the flower, a small net positive charge was detected in the stem, even before the bee landed. Once the bee landed, the whole plant became positively charged, and this positive charge remained on the plant after the bee flew away. By creating artificial flowers with various charge values, experimenters found that bees can distinguish between charged and uncharged flowers and may use the positive electric charge left by a previous bee as a cue indicating whether a plant has already been visited (in which case, little pollen may remain).

21.98 Consider a bee with the mean electric charge found in the experiment. This charge represents roughly how many missing electrons? (a) 1.9×10^8; (b) 3.0×10^8; (c) 1.9×10^{18}; (d) 3.0×10^{18}.

21.99 What is the best explanation for the observation that the electric charge on the stem became positive as the charged bee approached (before it landed)? (a) Because air is a good conductor, the positive charge on the bee's surface flowed through the air from bee to plant. (b) Because the earth is a reservoir of large amounts of charge, positive ions were drawn up the stem from the ground toward the charged bee. (c) The plant became electrically polarized as the charged bee approached. (d) Bees that had visited the plant earlier deposited a positive charge on the stem.

21.100 After one bee left a flower with a positive charge, that bee flew away and another bee with the same amount of positive charge flew close to the plant. Which diagram in **Fig. P21.100** (next page) best represents the electric field lines between the bee and the flower?

Figure **P21.100**



21.101 In a follow-up experiment, a charge of +40 pC was placed at the center of an artificial flower at the end of a 30-cm-long stem. Bees were observed to approach no closer than 15 cm from the center of this flower before they flew away. This observation suggests that the smallest external electric field to which bees may be sensitive is closest to which of these values? (a) 2.4 N/C; (b) 16 N/C; (c) 2.7×10^{-10} N/C; (d) 4.8×10^{-10} N/C.

ANSWERS

Chapter Opening Question ?

(ii) Water molecules have a permanent electric dipole moment: One end of the molecule has a positive charge and the other end has a negative charge. These ends attract negative and positive ions, respectively, holding the ions apart in solution. Water is less effective as a solvent for materials whose molecules do not ionize (called *nonionic* substances), such as oils.

Key Example VARIATION Problems

VP21.4.1 $(-98.9\ \mu\text{N})\hat{\imath}$
VP21.4.2 (a) 40.5 μN in the $+x$-direction (b) 13.6 μN in the $+x$-direction (c) +4.82 nC
VP21.4.3 3.45 μN in the $+y$-direction
VP21.4.4 $F_x = +0.452\ \mu\text{N}$, $F_y = +0.329\ \mu\text{N}$
VP21.10.1 (a) $E_x = 0$, $E_y = -8.09 \times 10^3$ N/C (b) $E_x = 0$, $E_y = +6.18 \times 10^2$ N/C (c) $E_x = -8.06 \times 10^2$ N/C, $E_y = -3.18 \times 10^2$ N/C
VP21.10.2 (a) 4.05×10^4 N/C in the $+x$-direction (b) 2.70×10^4 N/C in the $+x$-direction (c) -1.20 nC
VP21.10.3 (a) $E = \dfrac{e}{4\pi\epsilon_0 a^2}\left(1 - \dfrac{1}{2\sqrt{2}}\right)$ (b) away from the proton
VP21.10.4 (a) $dq = (Q/L)dx$ (b) $dE_x = -\dfrac{dq}{4\pi\epsilon_0 x^2} = -\dfrac{Q\,dx}{4\pi\epsilon_0 L x^2}$, $dE_y = 0$ (c) $E_x = -\dfrac{Q}{8\pi\epsilon_0 L^2}$, $E_y = 0$
VP21.14.1 (a) 1.41×10^{-24} N·m (b) -1.18×10^{-24} J
VP21.14.2 7.06×10^{-21} C
VP21.14.3 1.92×10^{-30} C·m
VP21.14.4 (a) 2.19×10^{-10} m (b) 1.99×10^{-8} m

Bridging Problem

$E = 2kQ/\pi a^2$ in the $-y$-direction

? This child acquires an electric charge by touching the charged metal shell. The charged hairs on the child's head repel and stand out. What would happen if the child stood *inside* a large, charged metal shell? She would acquire (i) the same sign of charge as on the shell, and her hairs would stand out; (ii) the opposite sign of charge as on the shell, and her hairs would stand out; (iii) no charge, and her hairs would be relaxed; (iv) any of these, depending on the amount of charge on the shell.

22 Gauss's Law

LEARNING OUTCOMES

In this chapter, you'll learn...

22.1 How you can determine the amount of charge within a closed surface by examining the electric field on the surface.
22.2 What is meant by electric flux, and how to calculate it.
22.3 How Gauss's law relates the electric flux through a closed surface to the charge enclosed by the surface.
22.4 How to use Gauss's law to calculate the electric field due to a symmetric charge distribution.
22.5 Where the charge is located on a charged conductor.

You'll need to review...

21.4–21.6 Electric fields and their properties.

In physics, an important tool for simplifying problems is the *symmetry properties* of systems. Many physical systems have symmetry; for example, a cylindrical body doesn't look any different after you've rotated it around its axis, and a charged metal sphere looks just the same after you've turned it about any axis through its center.

In this chapter we'll use symmetry ideas along with a new principle, called *Gauss's law,* to simplify electric-field calculations. For example, the field of a straight-line or plane-sheet charge distribution, which we derived in Section 21.5 by using some fairly strenuous integrations, can be obtained in a few steps with the help of Gauss's law. But Gauss's law is more than just a way to make certain calculations easier. Indeed, it is a fundamental statement about the relationship between electric charges and electric fields. Among other things, Gauss's law can help us understand how electric charge distributes itself over conducting bodies.

Here's what Gauss's law is all about. Given any general distribution of charge, we surround it with an imaginary surface that encloses the charge. Then we look at the electric field at various points on this imaginary surface. Gauss's law is a relationship between the field at *all* the points on the surface and the total charge enclosed within the surface. This may sound like a rather indirect way of expressing things, but it turns out to be a tremendously useful relationship. In the next several chapters, we'll make frequent use of the insights that Gauss's law provides into the character of electric fields.

22.1 CHARGE AND ELECTRIC FLUX

The discussion of Gauss's law in this section is based on and inspired by the innovative ideas of Ruth W. Chabay and Bruce A. Sherwood in *Electric and Magnetic Interactions* (John Wiley & Sons, 1994).

In Chapter 21 we asked the question, "Given a charge distribution, what is the electric field produced by that distribution at a point P?" We saw that the answer could be found by representing the distribution as an assembly of point charges, each of which produces an electric field $\vec{E}$ given by Eq. (21.7). The total field at P is then the vector sum of the fields due to all the point charges.

But there is an alternative relationship between charge distributions and electric fields. To discover this relationship, let's stand the question of Chapter 21 on its head and ask, "If the electric-field pattern is known in a given region, what can we determine about the charge distribution in that region?"

Here's an example. Consider the box shown in **Fig. 22.1a**, which may or may not contain electric charge. We'll imagine that the box is made of a material that has no effect on any electric fields; it's of the same breed as the massless rope and the frictionless incline. Better still, let the box represent an *imaginary* surface that may or may not enclose some charge. We'll refer to the box as a **closed surface** because it completely encloses a volume. How can you determine how much (if any) electric charge lies within the box?

Knowing that a charge distribution produces an electric field and that an electric field exerts a force on a test charge, you move a test charge q_0 around the vicinity of the box. By measuring the force $\vec{F}$ experienced by the test charge at different positions, you make a three-dimensional map of the electric field $\vec{E} = \vec{F}/q_0$ outside the box. In the case shown in Fig. 22.1b, the map turns out to be the same as that of the electric field produced by a positive point charge (Fig. 21.28a). From the details of the map, you can find the exact value of the point charge inside the box.

To determine the contents of the box, we actually need to measure $\vec{E}$ on only the *surface* of the box. In **Fig. 22.2a** there is a single *positive* point charge inside the box, and in Fig. 22.2b there are two such charges. The field patterns on the surfaces of the boxes are different in detail, but in each case the electric field points *out* of the box. Figures 22.2c and 22.2d show cases with one and two *negative* point charges, respectively, inside the box. Again, the details of $\vec{E}$ are different for the two cases, but the electric field points *into* each box.

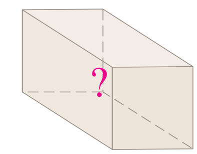
Electric Flux and Enclosed Charge

In Section 21.4 we mentioned the analogy between electric-field vectors and the velocity vectors of a fluid in motion. This analogy can be helpful, even though an electric field does not actually "flow." Using this analogy, in Figs. 22.2a and 22.2b, in which the electric-field vectors point out of the surface, we say that there is an *outward* **electric flux.** (The word "flux" comes from a Latin word meaning "flow.") In Figs. 22.2c and 22.2d the $\vec{E}$ vectors point into the surface, and the electric flux is *inward*.

Figure 22.2 suggests a simple relationship: Positive charge inside the box goes with an outward electric flux through the box's surface, and negative charge inside goes with an inward electric flux. What happens if there is *zero* charge inside the box? In **Fig. 22.3a** (next page) the box is empty and $\vec{E} = \mathbf{0}$ everywhere, so there is no electric flux into or out of the box. In Fig. 22.3b, one positive and one negative point charge of equal magnitude are enclosed within the box, so the *net* charge inside the box is zero. There is an electric field, but it "flows into" the box on half of its surface and "flows out of" the box on the other half. Hence there is no *net* electric flux into or out of the box.

Figure 22.1 How can you measure the charge inside a box without opening it?

(a) A box containing an unknown amount of charge



(b) Using a test charge outside the box to probe the amount of charge inside the box

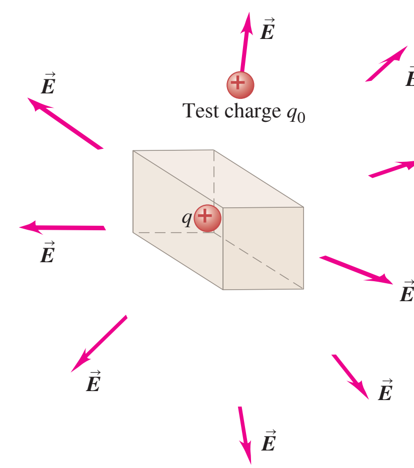


Figure 22.2 The electric field on the surface of boxes containing (a) a single positive point charge, (b) two positive point charges, (c) a single negative point charge, or (d) two negative point charges.

(a) Positive charge inside box, outward flux

(b) Positive charges inside box, outward flux

(c) Negative charge inside box, inward flux

(d) Negative charges inside box, inward flux

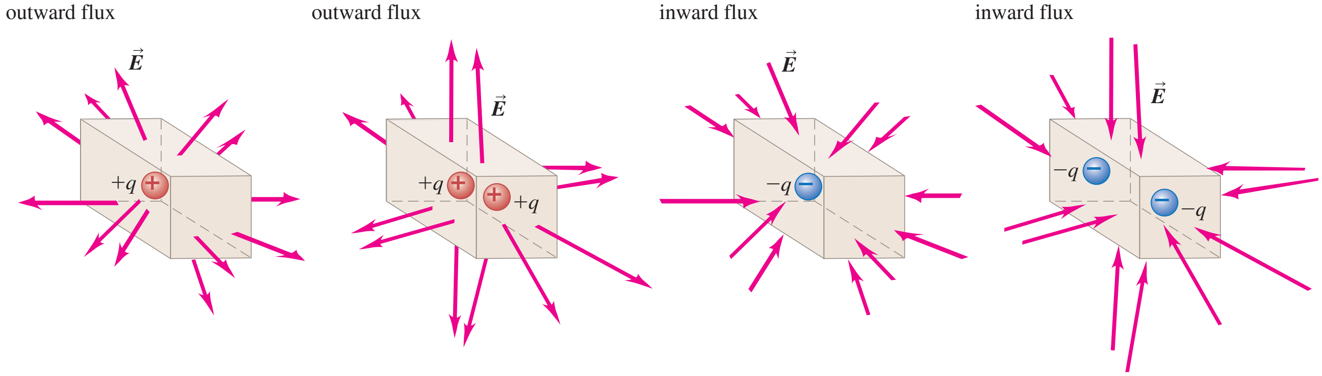


Figure **22.3** Three cases in which there is zero *net* charge inside a box and no net electric flux through the surface of the box. (a) An empty box with $\vec{E} = \mathbf{0}$. (b) A box containing one positive and one equal-magnitude negative point charge. (c) An empty box immersed in a uniform electric field.

(a) No charge inside box, zero flux

(b) Zero *net* charge inside box, inward flux cancels outward flux.

(c) No charge inside box, inward flux cancels outward flux.

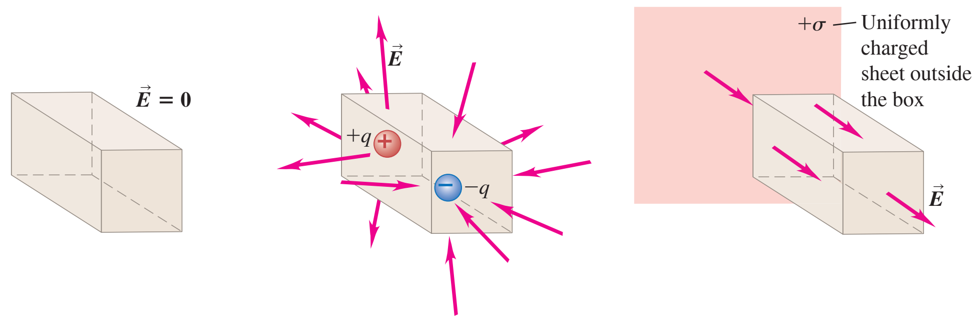
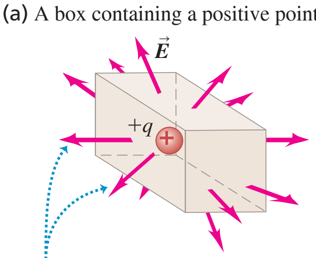
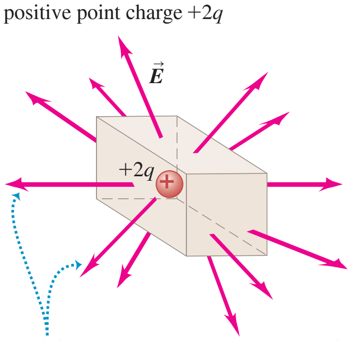


Figure **22.4** Three boxes, each of which encloses a positive point charge.

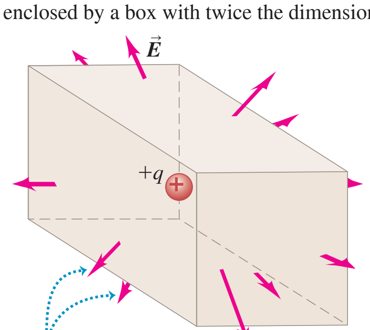
(a) A box containing a positive point charge $+q$



(b) The same box as in (a), but containing a positive point charge $+2q$



(c) The same positive point charge $+q$, but enclosed by a box with twice the dimensions



The box is again empty in Fig. 22.3c. However, there is charge present *outside* the box; the box has been placed with one end parallel to a uniformly charged infinite sheet, which produces a uniform electric field perpendicular to the sheet (see Example 21.11 of Section 21.5). On one end of the box, $\vec{E}$ points into the box; on the opposite end, $\vec{E}$ points out of the box; and on the sides, $\vec{E}$ is parallel to the surface and so points neither into nor out of the box. As in Fig. 22.3b, the inward electric flux on one part of the box exactly compensates for the outward electric flux on the other part. So in all of the cases shown in Fig. 22.3, there is no *net* electric flux through the surface of the box, and no *net* charge is enclosed in the box.

Figures 22.2 and 22.3 demonstrate a connection between the *sign* (positive, negative, or zero) of the *net* charge enclosed by a closed surface and the sense (outward, inward, or none) of the net electric flux through the surface. There is also a connection between the *magnitude* of the net charge inside the closed surface and the *strength* of the net "flow" of $\vec{E}$ over the surface. In both **Figs. 22.4a** and 22.4b there is a single point charge inside the box, but in Fig. 22.4b the magnitude of the charge is twice as great, and so $\vec{E}$ is everywhere twice as great in magnitude as in Fig. 22.4a. If we keep in mind the fluid-flow analogy, this means that the net outward electric flux is also twice as great in Fig. 22.4b as in Fig. 22.4a. This suggests that the net electric flux through the surface of the box is *directly proportional* to the magnitude of the net charge enclosed by the box.

This conclusion is independent of the size of the box. In Fig. 22.4c the point charge $+q$ is enclosed by a box with twice the linear dimensions of the box in Fig. 22.4a. The magnitude of the electric field of a point charge decreases with distance according to $1/r^2$, so the average magnitude of $\vec{E}$ on each face of the large box in Fig. 22.4c is just $\frac{1}{4}$ of the average magnitude on the corresponding face in Fig. 22.4a. But each face of the large box has exactly four times the area of the corresponding face of the small box. Hence the outward electric flux is the *same* for the two boxes if we *define* electric flux as follows: For each face of the box, take the product of the average perpendicular component of $\vec{E}$ and the area of that face; then add up the results from all faces of the box. With this definition the net electric flux due to a single point charge inside the box is independent of the size of the box and depends only on the net charge inside the box.

To summarize, for the special cases of a closed surface in the shape of a rectangular box and charge distributions made up of point charges or infinite charged sheets, we have found:

1. Whether there is a net outward or inward electric flux through a closed surface depends on the sign of the enclosed charge.
2. Charges *outside* the surface do not give a net electric flux through the surface.
3. The net electric flux is directly proportional to the net amount of charge enclosed within the surface but is otherwise independent of the size of the closed surface.

These observations are a qualitative statement of *Gauss's law*.

Do these observations hold true for other kinds of charge distributions and for closed surfaces of arbitrary shape? The answer to these questions will prove to be yes. But to explain why this is so, we need a precise mathematical statement of what we mean by electric flux. We develop this in the next section.

TEST YOUR UNDERSTANDING OF SECTION 22.1 If all of the dimensions of the box in Fig. 22.2a are increased by a factor of 3, how will the electric flux through the box change? (i) The flux will be $3^2 = 9$ times greater; (ii) the flux will be 3 times greater; (iii) the flux will be unchanged; (iv) the flux will be $\frac{1}{3}$ as great; (v) the flux will be $\left(\frac{1}{3}\right)^2 = \frac{1}{9}$ as great; (vi) not enough information is given to decide.

ANSWER

(iii) Each part of the surface of the box will be three times farther from the charge $+q$, so the electric field will be $\left(\frac{1}{3}\right)^2 = \frac{1}{9}$ as strong. But the area of the box will increase by a factor of $3^2 = 9$. Hence the electric flux will be multiplied by a factor of $\left(\frac{1}{9}\right)(9) = 1$. In other words, the flux will be unchanged.

22.2 CALCULATING ELECTRIC FLUX

In the preceding section we introduced the concept of *electric flux*. We used this to give a rough qualitative statement of Gauss's law: The net electric flux through a closed surface is directly proportional to the net charge inside that surface. To be able to make full use of this law, we need to know how to *calculate* electric flux. To do this, let's again make use of the analogy between an electric field $\vec{E}$ and the field of velocity vectors $\vec{v}$ in a flowing fluid. (Again, keep in mind that this is only an analogy; an electric field is *not* a flow.)

Flux: Fluid-Flow Analogy

Figure 22.5 shows a fluid flowing steadily from left to right. Let's examine the volume flow rate dV/dt (in, say, cubic meters per second) through the wire rectangle with area A. When the area is perpendicular to the flow velocity $\vec{v}$ (Fig. 22.5a) and the flow velocity is the same at all points in the fluid, the volume flow rate dV/dt is the area A multiplied by the flow speed v:

$$\frac{dV}{dt} = vA$$

When the rectangle is tilted at an angle ϕ (Fig. 22.5b) so that its face is not perpendicular to $\vec{v}$, the area that counts is the silhouette area that we see when we look in the direction of $\vec{v}$. This area, which is outlined in red and labeled $A_\perp$ in Fig. 22.5b, is the *projection* of the area A onto a surface perpendicular to $\vec{v}$. Two sides of the projected rectangle have the same length as the original one, but the other two are foreshortened by a factor of $\cos\phi$, so the projected area $A_\perp$ is equal to $A\cos\phi$. Then the volume flow rate through A is

$$\frac{dV}{dt} = vA\cos\phi$$

If $\phi = 90°$, $dV/dt = 0$; the wire rectangle is edge-on to the flow, and no fluid passes through the rectangle.

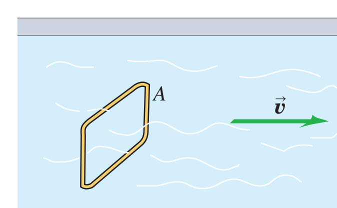
Also, $v\cos\phi$ is the component of the vector $\vec{v}$ perpendicular to the plane of the area A. Calling this component $v_\perp$, we can rewrite the volume flow rate as

$$\frac{dV}{dt} = v_\perp A$$

We can express the volume flow rate more compactly by using the concept of *vector area* $\vec{A}$, a vector quantity with magnitude A and a direction perpendicular to the plane of the area we are describing. The vector area $\vec{A}$ describes both the size of an area and its

Figure **22.5** The volume flow rate of fluid through the wire rectangle (a) is vA when the area of the rectangle is perpendicular to $\vec{v}$ and (b) is $vA\cos\phi$ when the rectangle is tilted at an angle ϕ.

(a) A wire rectangle in a fluid



(b) The wire rectangle tilted by an angle ϕ

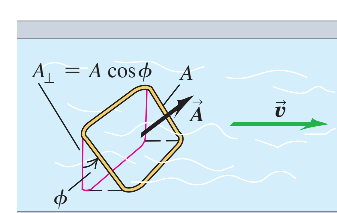
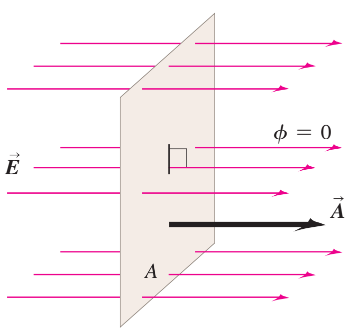


Figure 22.6 A flat surface in a uniform electric field. The electric flux Φ_E through the surface equals the scalar product of the electric field $\vec{E}$ and the area vector $\vec{A}$.

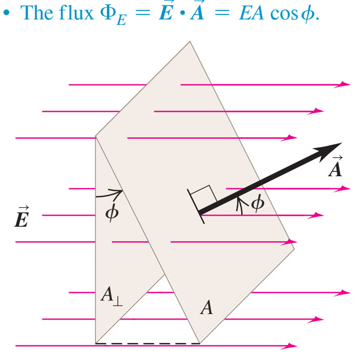
(a) Surface is face-on to electric field:

- $\vec{E}$ and $\vec{A}$ are parallel (the angle between $\vec{E}$ and $\vec{A}$ is $\phi = 0$).
- The flux $\Phi_E = \vec{E} \cdot \vec{A} = EA$.



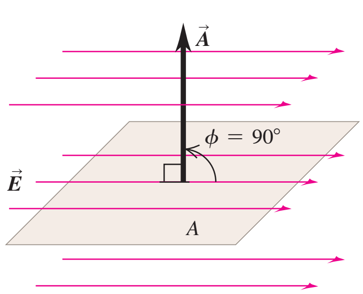
(b) Surface is tilted from a face-on orientation by an angle ϕ:

- The angle between $\vec{E}$ and $\vec{A}$ is ϕ.
- The flux $\Phi_E = \vec{E} \cdot \vec{A} = EA \cos\phi$.



(c) Surface is edge-on to electric field:

- $\vec{E}$ and $\vec{A}$ are perpendicular (the angle between $\vec{E}$ and $\vec{A}$ is $\phi = 90°$).
- The flux $\Phi_E = \vec{E} \cdot \vec{A} = EA \cos 90° = 0$.



orientation in space. In terms of $\vec{A}$, we can write the volume flow rate of fluid through the rectangle in Fig. 22.5b as a scalar (dot) product:

$$\frac{dV}{dt} = \vec{v} \cdot \vec{A}$$

Flux of a Uniform Electric Field

Using the analogy between electric field and fluid flow, we now define electric flux in the same way as we have just defined the volume flow rate of a fluid; we simply replace the fluid velocity $\vec{v}$ by the electric field $\vec{E}$. The symbol that we use for electric flux is Φ_E (the capital Greek letter phi; the subscript E is a reminder that this is *electric* flux). Consider first a flat area A perpendicular to a uniform electric field $\vec{E}$ (**Fig. 22.6a**). We define the electric flux through this area to be the product of the field magnitude E and the area A:

$$\Phi_E = EA$$

Roughly speaking, we can picture Φ_E in terms of the field lines passing through A. Increasing the area means that more lines of $\vec{E}$ pass through the area, increasing the flux; a stronger field means more closely spaced lines of $\vec{E}$ and therefore more lines per unit area, so again the flux increases.

If the area A is flat but not perpendicular to the field $\vec{E}$, then fewer field lines pass through it. In this case the area that counts is the silhouette area that we see when looking in the direction of $\vec{E}$. This is the area $A_\perp$ in Fig. 22.6b and is equal to $A\cos\phi$ (compare to Fig. 22.5b). We generalize our definition of electric flux for a uniform electric field to

$$\Phi_E = EA\cos\phi \qquad \text{(electric flux for uniform } \vec{E}\text{, flat surface)} \qquad (22.1)$$

Since $E\cos\phi$ is the component of $\vec{E}$ perpendicular to the area, we can rewrite Eq. (22.1) as

$$\Phi_E = E_\perp A \qquad \text{(electric flux for uniform } \vec{E}\text{, flat surface)} \qquad (22.2)$$

In terms of the vector area $\vec{A}$ perpendicular to the area, we can write the electric flux as the scalar product of $\vec{E}$ and $\vec{A}$:

$$\Phi_E = \vec{E} \cdot \vec{A} \qquad \text{(electric flux for uniform } \vec{E}\text{, flat surface)} \qquad (22.3)$$

Equations (22.1), (22.2), and (22.3) express the electric flux for a *flat* surface and a *uniform* electric field in different but equivalent ways. The SI unit for electric flux is $1\ \mathrm{N \cdot m^2/C}$. Note that if the area is edge-on to the field, $\vec{E}$ and $\vec{A}$ are perpendicular and the flux is zero (Fig. 22.6c).

We can represent the direction of a vector area $\vec{A}$ by using a *unit vector* $\hat{n}$ perpendicular to the area; $\hat{n}$ stands for "normal." Then

$$\vec{A} = A\hat{n} \qquad (22.4)$$

A surface has two sides, so there are two possible directions for $\hat{n}$ and $\vec{A}$. We must always specify which direction we choose. In Section 22.1 we related the charge inside a *closed* surface to the electric flux through the surface. With a closed surface we'll always choose the direction of $\hat{n}$ to be *outward,* and we'll speak of the flux *out of* a closed surface. Thus what we called "outward electric flux" in Section 22.1 corresponds to a *positive* value of Φ_E, and what we called "inward electric flux" corresponds to a *negative* value of Φ_E.

Flux of a Nonuniform Electric Field

What happens if the electric field $\vec{E}$ isn't uniform but varies from point to point over the area A? Or what if A is part of a curved surface? Then we divide A into many small elements dA, each of which has a unit vector $\hat{n}$ perpendicular to it and a vector area $d\vec{A} = \hat{n}\,dA$. We calculate the electric flux through each element and integrate the results to obtain the total flux:

$$\underset{\text{through a surface}}{\overset{\text{Electric flux}}{\Phi_E}} = \int \underset{\text{Angle between }\vec{E}\text{ and normal to surface}}{\overset{\text{Magnitude of electric field }\vec{E}}{E}\cos\phi}\, \underbrace{dA}_{\text{Element of surface area}} = \int \overset{\text{Component of }\vec{E}\text{ perpendicular to surface}}{E_{\perp}}\, dA = \int \vec{E}\cdot \underbrace{d\vec{A}}_{\text{Vector element of surface area}} \qquad (22.5)$$

We call this integral the **surface integral** of the component $E_{\perp}$ over the area, or the surface integral of $\vec{E}\cdot d\vec{A}$. In specific problems, one form of the integral is sometimes more convenient than another. Example 22.3 at the end of this section illustrates the use of Eq. (22.5).

In Eq. (22.5) the electric flux $\int E_{\perp}\,dA$ is equal to the *average* value of the perpendicular component of the electric field, multiplied by the area of the surface. This is the same definition of electric flux that we were led to in Section 22.1, now expressed more mathematically. In the next section we'll see the connection between the total electric flux through *any* closed surface, no matter what its shape, and the amount of charge enclosed within that surface.

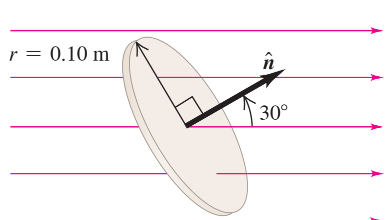
BIO APPLICATION Flux Through a Basking Shark's Mouth Unlike aggressive carnivorous sharks such as great whites, a basking shark feeds passively on plankton in the water that passes through the shark's gills as it swims. To survive on these tiny organisms requires a huge flux of water through a basking shark's immense mouth, which can be up to a meter across. The water flux—the product of the shark's speed through the water and the area of its mouth—can be up to 0.5 m^3/s (500 liters per second, or almost 5×10^5 gallons per hour). In a similar way, the flux of electric field through a surface depends on the magnitude of the field and the area of the surface (as well as the relative orientation of the field and surface).

EXAMPLE 22.1 Electric flux through a disk

A disk of radius 0.10 m is oriented with its normal unit vector $\hat{n}$ at 30° to a uniform electric field $\vec{E}$ of magnitude 2.0×10^3 N/C (**Fig. 22.7**). (Since this isn't a closed surface, it has no "inside" or "outside." That's why we have to specify the direction of $\hat{n}$ in the figure.) (a) What is the electric flux through the disk? (b) What is the flux through the disk if it is turned so that $\hat{n}$ is perpendicular to $\vec{E}$? (c) What is the flux through the disk if $\hat{n}$ is parallel to $\vec{E}$?

IDENTIFY and SET UP This problem is about a flat surface in a uniform electric field, so we can apply the ideas of this section. We calculate the electric flux from Eq. (22.1).

Figure 22.7 The electric flux Φ_E through a disk depends on the angle between its normal $\hat{n}$ and the electric field $\vec{E}$.



EXECUTE (a) The area is $A = \pi(0.10\text{ m})^2 = 0.0314\text{ m}^2$ and the angle between $\vec{E}$ and $\vec{A} = A\hat{n}$ is $\phi = 30°$, so from Eq. (22.1),

$$\Phi_E = EA\cos\phi = (2.0\times10^3\text{ N/C})(0.0314\text{ m}^2)(\cos 30°)$$
$$= 54\text{ N}\cdot\text{m}^2/\text{C}$$

(b) The normal to the disk is now perpendicular to $\vec{E}$, so $\phi = 90°$, $\cos\phi = 0$, and $\Phi_E = 0$.

(c) The normal to the disk is parallel to $\vec{E}$, so $\phi = 0$ and $\cos\phi = 1$:

$$\Phi_E = EA\cos\phi = (2.0\times10^3\text{ N/C})(0.0314\text{ m}^2)(1)$$
$$= 63\text{ N}\cdot\text{m}^2/\text{C}$$

EVALUATE Our answer to part (b) is smaller than that to part (a), which is in turn smaller than that to part (c). Is all this as it should be?

KEYCONCEPT To find the electric flux of a uniform electric field $\vec{E}$ through a flat surface, multiply the area of the surface by the component of $\vec{E}$ normal (perpendicular) to the surface.

EXAMPLE 22.2 Electric flux through a cube

WITH VARIATION PROBLEMS

An imaginary cubical surface of side L is in a region of uniform electric field $\vec{E}$. Find the electric flux through each face of the cube and the total flux through the cube when (a) it is oriented with two of its faces perpendicular to $\vec{E}$ (**Fig. 22.8a**) and (b) the cube is turned by an angle θ about a vertical axis (Fig. 22.8b).

IDENTIFY and SET UP Since $\vec{E}$ is uniform and each of the six faces of the cube is flat, we find the flux Φ_{Ei} through each face from Eqs. (22.3) and (22.4). The total flux through the cube is the sum of the six individual fluxes.

EXECUTE (a) Figure 22.8a shows the unit vectors $\hat{\boldsymbol{n}}_1$ through $\hat{\boldsymbol{n}}_6$ for each face; each unit vector points *outward* from the cube's closed surface. The angle between $\vec{E}$ and $\hat{\boldsymbol{n}}_1$ is 180°, the angle between $\vec{E}$ and $\hat{\boldsymbol{n}}_2$ is 0°, and the angle between $\vec{E}$ and each of the other four unit vectors is 90°. Each face of the cube has area L^2, so the fluxes through the faces are

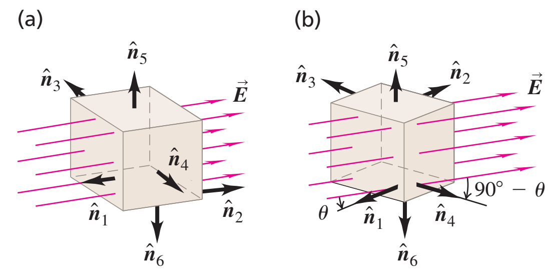
$$\Phi_{E1} = \vec{E} \cdot \hat{\boldsymbol{n}}_1 A = EL^2 \cos 180° = -EL^2$$
$$\Phi_{E2} = \vec{E} \cdot \hat{\boldsymbol{n}}_2 A = EL^2 \cos 0° = +EL^2$$
$$\Phi_{E3} = \Phi_{E4} = \Phi_{E5} = \Phi_{E6} = EL^2 \cos 90° = 0$$

The flux is negative on face 1, where $\vec{E}$ is directed into the cube, and positive on face 2, where $\vec{E}$ is directed out of the cube. The total flux through the cube is

$$\begin{aligned}\Phi_E &= \Phi_{E1} + \Phi_{E2} + \Phi_{E3} + \Phi_{E4} + \Phi_{E5} + \Phi_{E6} \\ &= -EL^2 + EL^2 + 0 + 0 + 0 + 0 = 0\end{aligned}$$

(b) The field $\vec{E}$ is directed into faces 1 and 3, so the fluxes through them are negative; $\vec{E}$ is directed out of faces 2 and 4, so the fluxes through them are positive. We find

Figure **22.8** Electric flux of a uniform field $\vec{E}$ through a cubical box of side L in two orientations.



$$\Phi_{E1} = \vec{E} \cdot \hat{\boldsymbol{n}}_1 A = EL^2 \cos(180° - \theta) = -EL^2 \cos\theta$$
$$\Phi_{E2} = \vec{E} \cdot \hat{\boldsymbol{n}}_2 A = +EL^2 \cos\theta$$
$$\Phi_{E3} = \vec{E} \cdot \hat{\boldsymbol{n}}_3 A = EL^2 \cos(90° + \theta) = -EL^2 \sin\theta$$
$$\Phi_{E4} = \vec{E} \cdot \hat{\boldsymbol{n}}_4 A = EL^2 \cos(90° - \theta) = +EL^2 \sin\theta$$
$$\Phi_{E5} = \Phi_{E6} = EL^2 \cos 90° = 0$$

The total flux $\Phi_E = \Phi_{E1} + \Phi_{E2} + \Phi_{E3} + \Phi_{E4} + \Phi_{E5} + \Phi_{E6}$ through the surface of the cube is again zero.

EVALUATE We came to the same conclusion in our discussion of Fig. 22.3c: There is zero net flux of a uniform electric field through a closed surface that contains no electric charge.

KEYCONCEPT When calculating the net electric flux through a closed three-dimensional surface, always choose the normal at every location on the surface to point outward. The net flux equals the sum of the fluxes through each part of the surface.

EXAMPLE 22.3 Electric flux through a sphere

WITH VARIATION PROBLEMS

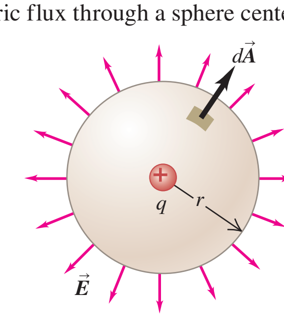
A point charge $q = +3.0\ \mu\text{C}$ is surrounded by an imaginary sphere of radius $r = 0.20$ m centered on the charge (**Fig. 22.9**). Find the resulting electric flux through the sphere.

IDENTIFY and SET UP The surface is not flat and the electric field is not uniform, so to calculate the electric flux (our target variable) we must use the general definition, Eq. (22.5). Because the sphere is centered on the point charge, at any point on the spherical surface, $\vec{E}$ is directed out of the sphere perpendicular to the surface. The positive direction for both $\hat{\boldsymbol{n}}$ and $E_\perp$ is outward, so $E_\perp = E$ and the flux through a surface element dA is $\vec{E} \cdot d\vec{A} = E\,dA$. This greatly simplifies the integral in Eq. (22.5).

EXECUTE We must evaluate the integral of Eq. (22.5), $\Phi_E = \int E\,dA$. At any point on the sphere of radius r the electric field has the same magnitude $E = q/4\pi\epsilon_0 r^2$. Hence E can be taken outside the integral, which becomes $\Phi_E = E\int dA = EA$, where A is the area of the spherical surface: $A = 4\pi r^2$. Hence the total flux through the sphere is

$$\begin{aligned}\Phi_E &= EA = \frac{q}{4\pi\epsilon_0 r^2} 4\pi r^2 = \frac{q}{\epsilon_0} \\ &= \frac{3.0 \times 10^{-6}\ \text{C}}{8.85 \times 10^{-12}\ \text{C}^2/\text{N} \cdot \text{m}^2} = 3.4 \times 10^5\ \text{N} \cdot \text{m}^2/\text{C}\end{aligned}$$

Figure **22.9** Electric flux through a sphere centered on a point charge.



EVALUATE The radius r of the sphere cancels out of the result for Φ_E. We would have obtained the same flux with a sphere of radius 2.0 m or 200 m. We came to essentially the same conclusion in our discussion of Fig. 22.4 in Section 22.1, where we considered rectangular closed surfaces of two different sizes enclosing a point charge. There we found that the flux of $\vec{E}$ was independent of the size of the surface; the same result holds true for a spherical surface. Indeed, the flux through *any* surface enclosing a single point charge is independent of the shape or size of the surface, as we'll soon see.

KEYCONCEPT If a closed surface is curved, calculating the net electric flux through the surface requires integrating the normal component of the electric field over the entire surface. It greatly simplifies the integration if the field is symmetrical and the surface used has the same symmetry.

TEST YOUR UNDERSTANDING OF SECTION 22.2 Rank the following surfaces in order from most positive to most negative electric flux: (i) a flat rectangular surface with vector area $\vec{A} = (6.0\text{ m}^2)\hat{\imath}$ in a uniform electric field $\vec{E} = (4.0\text{ N/C})\hat{\jmath}$; (ii) a flat circular surface with vector area $\vec{A} = (3.0\text{ m}^2)\hat{\jmath}$ in a uniform electric field $\vec{E} = (4.0\text{ N/C})\hat{\imath} + (2.0\text{ N/C})\hat{\jmath}$; (iii) a flat square surface with vector area $\vec{A} = (3.0\text{ m}^2)\hat{\imath} + (7.0\text{ m}^2)\hat{\jmath}$ in a uniform electric field $\vec{E} = (4.0\text{ N/C})\hat{\imath} - (2.0\text{ N/C})\hat{\jmath}$; (iv) a flat oval surface with vector area $\vec{A} = (3.0\text{ m}^2)\hat{\imath} - (7.0\text{ m}^2)\hat{\jmath}$ in a uniform electric field $\vec{E} = (4.0\text{ N/C})\hat{\imath} - (2.0\text{ N/C})\hat{\jmath}$.

ANSWER

(iv), (ii), (i), (iii) In each case the electric field is uniform, so the flux is $\Phi_E = \vec{E} \cdot \vec{A}$. We use the relationships for the scalar products of unit vectors: $\hat{\imath} \cdot \hat{\imath} = \hat{\jmath} \cdot \hat{\jmath} = 1$, $\hat{\imath} \cdot \hat{\jmath} = 0$. In case (i) we have $\Phi_E = (4.0\text{ N/C})(6.0\text{ m}^2)\hat{\imath} \cdot \hat{\jmath} = 0$ (the electric field and vector area are perpendicular, so there is zero flux). In case (ii) we have $\Phi_E = [(4.0\text{ N/C})\hat{\imath} + (2.0\text{ N/C})\hat{\jmath}] \cdot (3.0\text{ m}^2)\hat{\jmath} = (2.0\text{ N/C})(3.0\text{ m}^2) = 6.0\text{ N} \cdot \text{m}^2/\text{C}$. Similarly, in case (iii) we have $\Phi_E = [(4.0\text{ N/C})\hat{\imath} - (2.0\text{ N/C})\hat{\jmath}] \cdot [(3.0\text{ m}^2)\hat{\imath} + (7.0\text{ m}^2)\hat{\jmath}] = (4.0\text{ N/C})(3.0\text{ m}^2) - (2.0\text{ N/C})(7.0\text{ m}^2) = -2\text{ N} \cdot \text{m}^2/\text{C}$, and in case (iv) we have $\Phi_E = [(4.0\text{ N/C})\hat{\imath} - (2.0\text{ N/C})\hat{\jmath}] \cdot [(3.0\text{ m}^2)\hat{\imath} - (7.0\text{ m}^2)\hat{\jmath}] = (4.0\text{ N/C})(3.0\text{ m}^2) + (2.0\text{ N/C})(7.0\text{ m}^2) = 26\text{ N} \cdot \text{m}^2/\text{C}$.

22.3 GAUSS'S LAW

Gauss's law is an alternative to Coulomb's law. While completely equivalent to Coulomb's law, Gauss's law provides a different way to express the relationship between electric charge and electric field. It was formulated by Carl Friedrich Gauss (1777–1855), one of the greatest mathematicians of all time (**Fig. 22.10**).

Point Charge Inside a Spherical Surface

Gauss's law states that the total electric flux through any closed surface (a surface enclosing a definite volume) is proportional to the total (net) electric charge inside the surface. In Section 22.1 we observed this relationship qualitatively; now we'll develop it more rigorously. We'll start with the field of a single positive point charge q. The field lines radiate out equally in all directions. We place this charge at the center of an imaginary spherical surface with radius R. The magnitude E of the electric field at every point on the surface is given by

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{R^2}$$

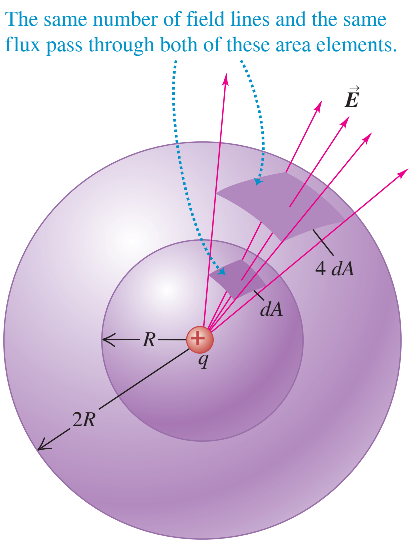
At each point on the surface, $\vec{E}$ is perpendicular to the surface, and its magnitude is the same at every point, as in Example 22.3 (Section 22.2). The total electric flux is the product of the field magnitude E and the total area $A = 4\pi R^2$ of the sphere:

$$\Phi_E = EA = \frac{1}{4\pi\epsilon_0}\frac{q}{R^2}(4\pi R^2) = \frac{q}{\epsilon_0} \tag{22.6}$$

The flux is independent of the radius R of the sphere. It depends on only the charge q enclosed by the sphere.

Figure **22.10** Carl Friedrich Gauss helped develop several branches of mathematics, including differential geometry, real analysis, and number theory. The "bell curve" of statistics is one of his inventions. Gauss also made state-of-the-art investigations of the earth's magnetism and calculated the orbit of the first asteroid to be discovered.

Figure **22.11** Projection of an element of area dA of a sphere of radius R onto a concentric sphere of radius $2R$. The projection multiplies each linear dimension by 2, so the area element on the larger sphere is $4\,dA$.



We can also interpret this result in terms of field lines. **Figure 22.11** shows two spheres with radii R and $2R$ centered on the point charge q. Every field line that passes through the smaller sphere also passes through the larger sphere, so the total flux through each sphere is the same.

What is true of the entire sphere is also true of any portion of its surface. In Fig. 22.11 an area dA is outlined on the sphere of radius R and projected onto the sphere of radius $2R$ by drawing lines from the center through points on the boundary of dA. The area projected on the larger sphere is clearly $4\,dA$. But since the electric field due to a point charge is inversely proportional to r^2, the field magnitude is $\frac{1}{4}$ as great on the sphere of radius $2R$ as on the sphere of radius R. Hence the electric flux is the same for both areas and is independent of the radius of the sphere.

Point Charge Inside a Nonspherical Surface

We can extend this projection technique to nonspherical surfaces. Instead of a second sphere, let us surround the sphere of radius R by a surface of irregular shape, as in **Fig. 22.12a**. Consider a small element of area dA on the irregular surface; we note that this area is *larger* than the corresponding element on a spherical surface at the same distance from q. If a normal to dA makes an angle ϕ with a radial line from q, two sides of the area projected onto the spherical surface are foreshortened by a factor $\cos\phi$ (Fig. 22.12b). The other two sides are unchanged. Thus the electric flux through the spherical surface element is equal to the flux $E\,dA\cos\phi$ through the corresponding irregular surface element.

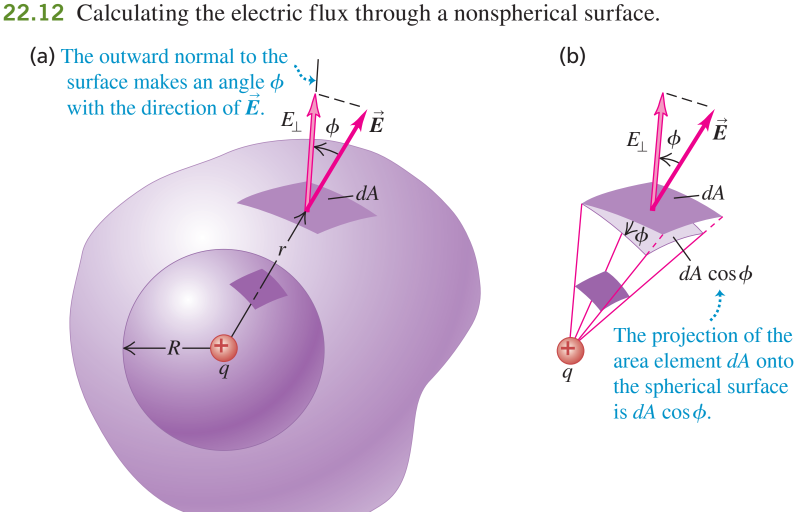
We can divide the entire irregular surface into elements dA, compute the electric flux $E\,dA\cos\phi$ for each, and sum the results by integrating, as in Eq. (22.5). Each of the area elements projects onto a corresponding spherical surface element. Thus the *total* electric flux through the irregular surface, given by any of the forms of Eq. (22.5), must be the same as the total flux through a sphere, which Eq. (22.6) shows is equal to q/ϵ_0. Thus, for the irregular surface,

$$\Phi_E = \oint \vec{E} \cdot d\vec{A} = \frac{q}{\epsilon_0} \tag{22.7}$$

Equation (22.7) holds for a surface of *any* shape or size, provided only that it is a *closed* surface enclosing the charge q. The circle on the integral sign reminds us that the integral is always taken over a *closed* surface.

The area elements $d\vec{A}$ and the corresponding unit vectors $\hat{n}$ always point *out of* the volume enclosed by the surface. The electric flux is then positive in areas where the electric field points out of the surface and negative where it points inward. Also, $E_\perp$ is positive at

Figure **22.12** Calculating the electric flux through a nonspherical surface.



points where $\vec{E}$ points out of the surface and negative at points where $\vec{E}$ points into the surface.

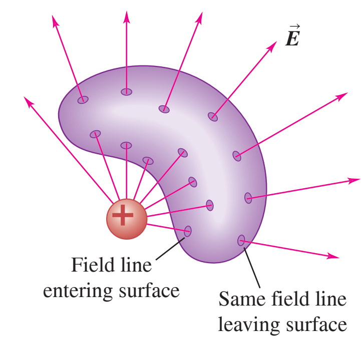
If the point charge in Fig. 22.12 is negative, the $\vec{E}$ field is directed radially *inward;* the angle ϕ is then greater than 90°, its cosine is negative, and the integral in Eq. (22.7) is negative. But since q is also negative, Eq. (22.7) holds.

For a closed surface enclosing *no* charge,

$$\Phi_E = \oint \vec{E} \cdot d\vec{A} = 0$$

This is the mathematical statement that when a region contains no charge, any field lines caused by charges *outside* the region that enter on one side must leave again on the other side. (In Section 22.1 we came to the same conclusion by considering the special case of a rectangular box in a uniform field.) **Figure 22.13** illustrates this point. *Electric field lines can begin or end inside a region of space only when there is charge in that region.*

Figure **22.13** A point charge *outside* a closed surface that encloses no charge. If an electric field line from the external charge enters the surface at one point, it must leave at another.



General Form of Gauss's Law

Now comes the final step in obtaining the general form of Gauss's law. Suppose the surface encloses not just one point charge q but several charges $q_1, q_2, q_3, \ldots$. The total (resultant) electric field $\vec{E}$ at any point is the vector sum of the $\vec{E}$ fields of the individual charges. Let Q_{encl} be the *total* charge enclosed by the surface: $Q_{\text{encl}} = q_1 + q_2 + q_3 + \cdots$. Also let $\vec{E}$ be the *total* field at the position of the surface area element $d\vec{A}$, and let $E_\perp$ be its component perpendicular to the plane of that element (that is, parallel to $d\vec{A}$). Then we can write an equation like Eq. (22.7) for each charge and its corresponding field and add the results. When we do, we obtain the general statement of Gauss's law:

Gauss's law:

$$\Phi_E = \oint \vec{E} \cdot d\vec{A} = \frac{Q_{\text{encl}}}{\epsilon_0} \tag{22.8}$$

Total charge enclosed by surface

Electric flux through a closed surface of area A = surface integral of $\vec{E}$

Electric constant

The total electric flux through a closed surface is equal to the total (net) electric charge inside the surface, divided by ϵ_0.

CAUTION **Gaussian surfaces are imaginary** Remember that the closed surface in Gauss's law is *imaginary;* there need not be any material object at the position of the surface. We often refer to a closed surface used in Gauss's law as a **Gaussian surface.**

Using the definition of Q_{encl} and the various ways to express electric flux given in Eq. (22.5), we can express Gauss's law in the following equivalent forms:

Various forms of Gauss's law:

$$\Phi_E = \oint E\cos\phi\, dA = \oint E_\perp\, dA = \oint \vec{E} \cdot d\vec{A} = \frac{Q_{\text{encl}}}{\epsilon_0} \tag{22.9}$$

Magnitude of electric field $\vec{E}$

Component of $\vec{E}$ perpendicular to surface

Total charge enclosed by surface

Electric flux through a closed surface

Angle between $\vec{E}$ and normal to surface

Element of surface area

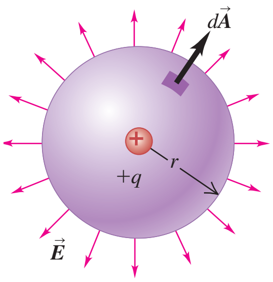
Vector element of surface area

Electric constant

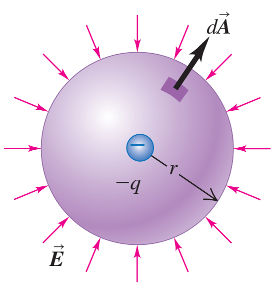
As in Eq. (22.5), the various forms of the integral all express the same thing, the total electric flux through the Gaussian surface, in different terms. One form is sometimes more convenient than another.

As an example, **Fig. 22.14a** (next page) shows a spherical Gaussian surface of radius r around a positive point charge $+q$. The electric field points out of the Gaussian surface, so at every point on the surface $\vec{E}$ is in the same direction as $d\vec{A}$, $\phi = 0$, and $E_\perp$ is equal to the field magnitude $E = q/4\pi\epsilon_0 r^2$. Since E is the same at all points on the surface, Figure **22.14** Spherical Gaussian surfaces around (a) a positive point charge and (b) a negative point charge.

(a) Gaussian surface around positive charge: positive (outward) flux



(b) Gaussian surface around negative charge: negative (inward) flux



we can take it outside the integral in Eq. (22.9). Then the remaining integral is $\int dA = A = 4\pi r^2$, the area of the sphere. Hence Eq. (22.9) becomes

$$\Phi_E = \oint E_\perp \, dA = \oint \left(\frac{q}{4\pi\epsilon_0 r^2} \right) dA = \frac{q}{4\pi\epsilon_0 r^2} \oint dA = \frac{q}{4\pi\epsilon_0 r^2} 4\pi r^2 = \frac{q}{\epsilon_0}$$

The enclosed charge Q_{encl} is just the charge $+q$, so this agrees with Gauss's law. If the Gaussian surface encloses a *negative* point charge as in Fig. 22.14b, then $\vec{E}$ points *into* the surface at each point in the direction opposite $d\vec{A}$. Then $\phi = 180°$ and $E_\perp$ is equal to the negative of the field magnitude: $E_\perp = -E = -|-q|/4\pi\epsilon_0 r^2 = -q/4\pi\epsilon_0 r^2$. Equation (22.9) then becomes

$$\Phi_E = \oint E_\perp \, dA = \oint \left(\frac{-q}{4\pi\epsilon_0 r^2} \right) dA = \frac{-q}{4\pi\epsilon_0 r^2} \oint dA = \frac{-q}{4\pi\epsilon_0 r^2} 4\pi r^2 = \frac{-q}{\epsilon_0}$$

This again agrees with Gauss's law because the enclosed charge in Fig. 22.14b is $Q_{\text{encl}} = -q$.

In Eqs. (22.8) and (22.9), Q_{encl} is always the algebraic sum of all the positive and negative charges enclosed by the Gaussian surface, and $\vec{E}$ is the *total* field at each point on the surface. Also note that in general, this field is caused partly by charges inside the surface and partly by charges outside. But as Fig. 22.13 shows, the outside charges do *not* contribute to the total (net) flux through the surface. So Eqs. (22.8) and (22.9) are correct even when there are charges outside the surface that contribute to the electric field at the surface. When $Q_{\text{encl}} = 0$, the total flux through the Gaussian surface must be zero, even though some areas may have positive flux and others may have negative flux (see Fig. 22.3b).

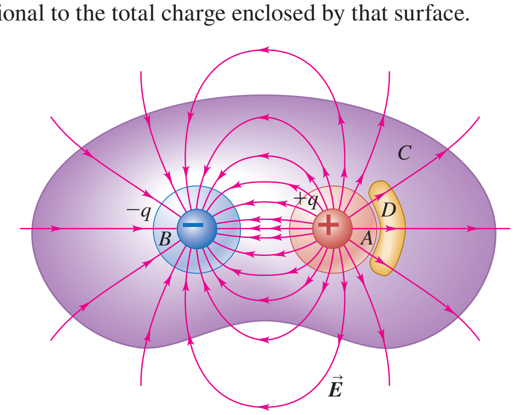
Gauss's law is the definitive answer to the question we posed at the beginning of Section 22.1: "If the electric-field pattern is known in a given region, what can we determine about the charge distribution in that region?" It provides a relationship between the electric field on a closed surface and the charge distribution within that surface. But in some cases we can use Gauss's law to answer the reverse question: "If the charge distribution is known, what can we determine about the electric field that the charge distribution produces?" Gauss's law may seem like an unappealing way to address this question, since it may look as though evaluating the integral in Eq. (22.8) is a hopeless task. Sometimes it is, but other times it is surprisingly easy. Here's an example in which *no* integration is involved at all; we'll work out several more examples in the next section.

CONCEPTUAL EXAMPLE 22.4 Electric flux and enclosed charge

WITH VARIATION PROBLEMS

Figure 22.15 shows the field produced by two point charges $+q$ and $-q$ (an electric dipole). Find the electric flux through each of the closed surfaces A, B, C, and D.

Figure **22.15** The net number of field lines leaving a closed surface is proportional to the total charge enclosed by that surface.



SOLUTION Gauss's law, Eq. (22.8), says that the total electric flux through a closed surface is equal to the total enclosed charge divided by ϵ_0. In Fig. 22.15, surface A (shown in red) encloses the positive charge, so $Q_{\text{encl}} = +q$; surface B (in blue) encloses the negative charge, so $Q_{\text{encl}} = -q$; surface C (in purple) encloses *both* charges, so $Q_{\text{encl}} = +q + (-q) = 0$; and surface D (in yellow) encloses no charges, so $Q_{\text{encl}} = 0$. Hence, without having to do any integration, we have $\Phi_{EA} = +q/\epsilon_0$, $\Phi_{EB} = -q/\epsilon_0$, and $\Phi_{EC} = \Phi_{ED} = 0$. These results depend only on the charges enclosed within each Gaussian surface, not on the precise shapes of the surfaces.

We can draw similar conclusions by examining the electric field lines. All the field lines that cross surface A are directed out of the surface, so the flux through A must be positive. Similarly, the flux through B must be negative since all of the field lines that cross that surface point inward. For both surface C and surface D, there are as many field lines pointing into the surface as there are field lines pointing outward, so the flux through each of these surfaces is zero.

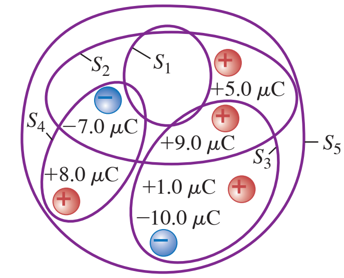
KEYCONCEPT The *net* electric flux Φ_E through a closed surface is always proportional to the *net* amount of charge q that the surface encloses. If q is zero, Φ_E is zero. These statements (Gauss's law) are true no matter what the shape of the surface.

TEST YOUR UNDERSTANDING OF SECTION 22.3 **Figure 22.16** shows six point charges that all lie in the same plane. Five Gaussian surfaces—S_1, S_2, S_3, S_4, and S_5— each enclose part of this plane, and Fig. 22.16 shows the intersection of each surface with the plane. Rank these five surfaces in order of the electric flux through them, from most positive to most negative.

ANSWER

❙ S_2, S_5, S_4, S_1 and S_3 (tie) Gauss's law tells us that the flux through a closed surface is proportional to the amount of charge enclosed within that surface. So an ordering of these surfaces by their fluxes is the same as an ordering by the amount of enclosed charge. Surface S_1 encloses no charge, surface S_2 encloses $9.0\ \mu C + 5.0\ \mu C + (-7.0\ \mu C) = 7.0\ \mu C$, surface S_3 encloses $9.0\ \mu C + 1.0\ \mu C + (-10.0\ \mu C) = 0$, surface S_4 encloses $8.0\ \mu C + (-7.0\ \mu C) = 1.0\ \mu C$, and surface S_5 encloses $8.0\ \mu C + (-7.0\ \mu C) + (-10.0\ \mu C) + (1.0\ \mu C) + (9.0\ \mu C) + (5.0\ \mu C) = 6.0\ \mu C$.

Figure **22.16** Five Gaussian surfaces and six point charges.



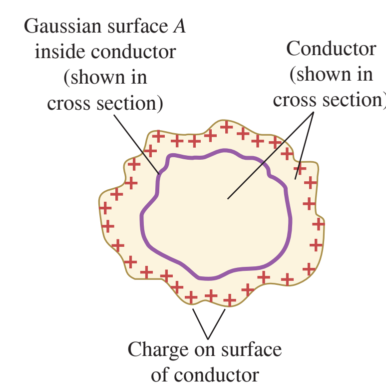
22.4 APPLICATIONS OF GAUSS'S LAW

Gauss's law is valid for *any* distribution of charges and for *any* closed surface. Gauss's law can be used in two ways. If we know the charge distribution, and if it has enough symmetry to let us evaluate the integral in Gauss's law, we can find the field. Or if we know the field, we can use Gauss's law to find the charge distribution, such as charges on conducting surfaces.

In this section we present examples of both kinds of applications. As you study them, watch for the role played by the symmetry properties of each system. We'll use Gauss's law to calculate the electric fields caused by several simple charge distributions; the results are collected in a table in the chapter summary.

In practical problems we often encounter situations in which we want to know the electric field caused by a charge distribution on a conductor. These calculations are aided by the following remarkable fact: *When excess charge is placed on a solid conductor and is at rest, it resides entirely on the surface, not in the interior of the material.* (By *excess* we mean charges other than the ions and free electrons that make up the neutral conductor.) Here's the proof. We know from Section 21.4 that in an electrostatic situation (with all charges at rest) the electric field $\vec{E}$ at every point in the interior of a conducting material is zero. If $\vec{E}$ were *not* zero, the excess charges would move. Suppose we construct a Gaussian surface inside the conductor, such as surface A in **Fig. 22.17**. Because $\vec{E} = \mathbf{0}$ everywhere on this surface, Gauss's law requires that the net charge inside the surface is zero. Now imagine shrinking the surface like a collapsing balloon until it encloses a region so small that we may consider it as a point P; then the charge at that point must be zero. We can do this anywhere inside the conductor, so *there can be no excess charge at any point within a solid conductor; any excess charge must reside on the conductor's surface.* (This result is for a *solid* conductor. In the next section we'll discuss what can happen if the conductor has cavities in its interior.) We'll make use of this fact frequently in the examples that follow.

Figure **22.17** Under electrostatic conditions (charges not in motion), any excess charge on a solid conductor resides entirely on the conductor's surface.



PROBLEM-SOLVING STRATEGY 22.1 Gauss's Law

IDENTIFY *the relevant concepts:* Gauss's law is most useful when the charge distribution has spherical, cylindrical, or planar symmetry. In these cases the symmetry determines the direction of $\vec{E}$. Then Gauss's law yields the magnitude of $\vec{E}$ if we are given the charge distribution, and vice versa. In either case, begin the analysis by asking the question: What is the symmetry?

SET UP *the problem* using the following steps:

1. List the known and unknown quantities and identify the target variable.
2. Select the appropriate closed, imaginary Gaussian surface. For spherical symmetry, use a concentric spherical surface. For cylindrical symmetry, use a coaxial cylindrical surface with flat ends perpendicular to the axis of symmetry (like a soup can). For planar symmetry, use a cylindrical surface (like a tuna can) with its flat ends parallel to the plane.

EXECUTE *the solution* as follows:

1. Determine the appropriate size and placement of your Gaussian surface. To evaluate the field magnitude at a particular point, the surface must include that point. It may help to place one end of a can-shaped surface within a conductor, where $\vec{E}$ and therefore Φ are zero, or to place its ends equidistant from a charged plane.
2. Evaluate the integral $\oint E_\perp\, dA$ in Eq. (22.9). In this equation $E_\perp$ is the perpendicular component of the *total* electric field at each point on the Gaussian surface. A well-chosen Gaussian surface should make integration trivial or unnecessary. If the surface comprises several separate surfaces, such as the sides and ends of a cylinder, the integral $\oint E_\perp\, dA$ over the entire closed surface is the sum of the integrals $\int E_\perp\, dA$ over the separate surfaces. Consider points 3–6 as you work.

Continued

3. If $\vec{E}$ is *perpendicular* (normal) at every point to a surface with area A, if it points *outward* from the interior of the surface, and if it has the same *magnitude* at every point on the surface, then $E_\perp = E =$ constant, and $\int E_\perp\, dA$ over that surface is equal to EA. (If $\vec{E}$ is inward, then $E_\perp = -E$ and $\int E_\perp\, dA = -EA$.) This should be the case for part or all of your Gaussian surface. If $\vec{E}$ is tangent to a surface at every point, then $E_\perp = 0$ and the integral over that surface is zero. This may be the case for parts of a cylindrical Gaussian surface. If $\vec{E} = \mathbf{0}$ at every point on a surface, the integral is zero.
4. Even when there is *no* charge within a Gaussian surface, the field at any given point on the surface is not necessarily zero. In that case, however, the total electric flux through the surface is always zero.
5. The flux integral $\oint E_\perp\, dA$ can be approximated as the difference between the numbers of electric lines of force leaving and entering the Gaussian surface. In this sense the flux gives the sign of the enclosed charge, but is only proportional to it; zero flux corresponds to zero enclosed charge.
6. Once you have evaluated $\oint E_\perp\, dA$, use Eq. (22.9) to solve for your target variable.

EVALUATE *your answer:* If your result is a *function* that describes how the magnitude of the electric field varies with position, ensure that it makes sense.

EXAMPLE 22.5 Field of a charged conducting sphere

WITH VARIATION PROBLEMS

We place a total positive charge q on a solid conducting sphere with radius R (**Fig. 22.18**). Find $\vec{E}$ at any point inside or outside the sphere.

IDENTIFY and SET UP As we discussed earlier in this section, all of the charge must be on the surface of the sphere. The charge is free to move on the conductor, and there is no preferred position on the surface; the charge is therefore distributed *uniformly* over the surface, and the system is spherically symmetric. To exploit this symmetry, we take as our Gaussian surface a sphere of radius r centered on the conductor. We can calculate the field inside or outside the conductor by taking $r < R$ or $r > R$, respectively. In either case, the point at which we want to calculate $\vec{E}$ lies on the Gaussian surface.

EXECUTE The spherical symmetry means that the direction of the electric field must be radial; that's because there is no preferred direction parallel to the surface, so $\vec{E}$ can have no component parallel to the surface. There is also no preferred orientation of the sphere, so the field magnitude E can depend only on the distance r from the center and must have the same value at all points on the Gaussian surface.

For $r > R$ the entire conductor is within the Gaussian surface, so the enclosed charge is q. The area of the Gaussian surface is $4\pi r^2$, and $\vec{E}$ is uniform over the surface and perpendicular to it at each point. The flux integral $\oint E_\perp\, dA$ is then just $E(4\pi r^2)$, and Eq. (22.8) gives

$$E(4\pi r^2) = \frac{q}{\epsilon_0} \quad \text{and}$$

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} \quad \text{(outside a charged conducting sphere)}$$

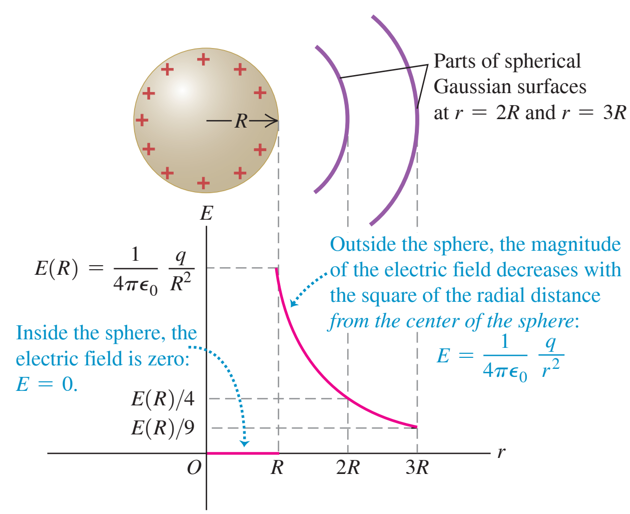
This expression is the same as that for a point charge; outside the charged sphere, its field is the same as though the entire charge were concentrated at its center. Just outside the surface of the sphere, where $r = R$,

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{R^2} \quad \text{(at the surface of a charged conducting sphere)}$$

CAUTION Flux can be positive or negative Remember that we have chosen the charge q to be *positive*. If the charge is negative, the electric field is radially *inward* instead of radially outward, and the electric flux through the Gaussian surface is negative. The electric-field magnitudes outside and at the surface of the sphere are given by the same expressions as above, except that q denotes the *magnitude* (absolute value) of the charge.

For $r < R$ we again have $E(4\pi r^2) = Q_{\text{encl}}/\epsilon_0$. But now our Gaussian surface (which lies entirely within the conductor) encloses *no* charge, so $Q_{\text{encl}} = 0$. The electric field inside the conductor is therefore zero.

Figure 22.18 Calculating the electric field of a conducting sphere with positive charge q. Outside the sphere, the field is the same as if all of the charge were concentrated at the center of the sphere.



EVALUATE We already knew that $\vec{E} = \mathbf{0}$ inside a solid conductor (whether spherical or not) when the charges are at rest. Figure 22.18 shows E as a function of the distance r from the center of the sphere. Note that in the limit as $R \to 0$, the sphere becomes a point charge; there is then only an "outside," and the field is everywhere given by $E = q/4\pi\epsilon_0 r^2$. Thus we have deduced Coulomb's law from Gauss's law. (In Section 22.3 we deduced Gauss's law from Coulomb's law; the two laws are equivalent.)

We can also use this method for a conducting spherical *shell* (a spherical conductor with a concentric spherical hole inside) if there is no charge inside the hole. We use a spherical Gaussian surface with radius r less than the radius of the hole. If there *were* a field inside the hole, it would have to be radial and spherically symmetric as before, so $E = Q_{\text{encl}}/4\pi\epsilon_0 r^2$. But now there is no enclosed charge, so $Q_{\text{encl}} = 0$ and $E = 0$ inside the hole.

Can you use this same technique to find the electric field in the region between a charged sphere and a concentric hollow conducting sphere that surrounds it?

KEYCONCEPT To apply Gauss's law to a charge distribution with spherical symmetry, use a spherical Gaussian surface. At all points on such a surface, the electric field has the same magnitude and is normal to the surface.

EXAMPLE 22.6 Field of a uniform line charge

WITH VARIATION PROBLEMS

Electric charge is distributed uniformly along an infinitely long, thin wire. The charge per unit length is λ (assumed positive). Find the electric field by using Gauss's law.

IDENTIFY and SET UP We found in Example 21.10 (Section 21.5) that the field $\vec{E}$ of a uniformly charged, infinite wire is radially outward if λ is positive and radially inward if λ is negative, and that the field magnitude E depends on only the radial distance from the wire. This suggests that we use a *cylindrical* Gaussian surface, of radius r and arbitrary length l, coaxial with the wire and with its ends perpendicular to the wire (**Fig. 22.19**).

EXECUTE The flux through the flat ends of our Gaussian surface is zero because the radial electric field is parallel to these ends, and so $\vec{E} \cdot \hat{n} = 0$. On the cylindrical part of our surface we have $\vec{E} \cdot \hat{n} = E_\perp = E$ everywhere. (If λ were negative, we would have $\vec{E} \cdot \hat{n} = E_\perp = -E$ everywhere.) The area of the cylindrical surface is $2\pi rl$, so the flux through it—and hence the total flux Φ_E through the Gaussian surface—is $EA = 2\pi rlE$. The total enclosed charge is $Q_{\text{encl}} = \lambda l$, and so from Gauss's law, Eq. (22.8),

$$\Phi_E = 2\pi rlE = \frac{\lambda l}{\epsilon_0} \quad \text{and}$$

$$E = \frac{1}{2\pi\epsilon_0}\frac{\lambda}{r} \qquad \text{(field of an infinite line of charge)}$$

We found this same result in Example 21.10 with *much* more effort.

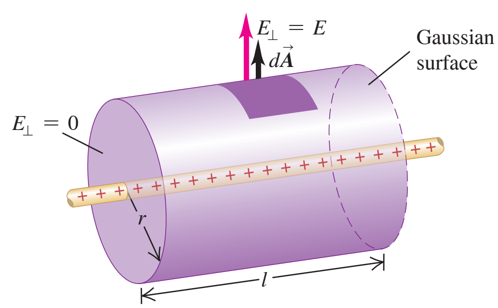
If λ is *negative*, $\vec{E}$ is directed radially inward, and in the above expression for E we must interpret λ as the absolute value of the charge per unit length.

EVALUATE We saw in Example 21.10 that the *entire* charge on the wire contributes to the field at any point, and yet we consider only that part of the charge $Q_{\text{encl}} = \lambda l$ within the Gaussian surface when we apply Gauss's law. There's nothing inconsistent here; it takes the entire charge to give the field the properties that allow us to calculate Φ_E so easily, and Gauss's law always applies to the enclosed charge only. If the wire is short, the symmetry of the infinite wire is lost, and E is not uniform over a coaxial, cylindrical Gaussian surface. Gauss's law then *cannot* be used to find Φ_E; we must solve the problem the hard way, as in Example 21.10.

We can use the Gaussian surface in Fig. 22.19 to show that the field outside a long, uniformly charged cylinder is the same as though all the charge were concentrated on a line along its axis (see Problem 22.39). We can also calculate the electric field in the space between a charged cylinder and a coaxial hollow conducting cylinder surrounding it (see Problem 22.37).

KEYCONCEPT To apply Gauss's law to a charge distribution with cylindrical symmetry, such as an infinite line or cylinder of charge, use a cylindrical Gaussian surface. At all points on the curved sides of the surface, the electric field has the same magnitude and is normal to the surface. On the flat ends of the cylindrical surface, the electric field is parallel to the surface, so the flux through the ends is zero.

Figure 22.19 A coaxial cylindrical Gaussian surface is used to find the electric field outside an infinitely long, charged wire.



EXAMPLE 22.7 Field of an infinite plane sheet of charge

WITH VARIATION PROBLEMS

Use Gauss's law to find the electric field caused by a thin, flat, infinite sheet with a uniform positive surface charge density σ.

IDENTIFY and SET UP In Example 21.11 (Section 21.5) we found that the field $\vec{E}$ of a uniformly charged infinite sheet is normal to the sheet, and that its magnitude is independent of the distance from the sheet. To take advantage of these symmetry properties, we use a cylindrical Gaussian surface with ends of area A and with its axis perpendicular to the sheet of charge (**Fig. 22.20**).

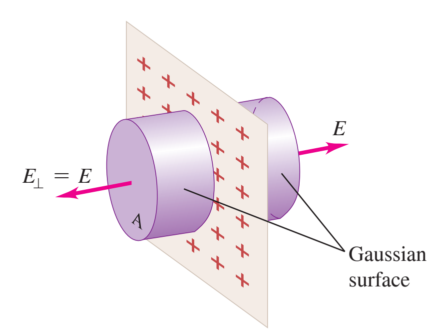
EXECUTE The flux through the cylindrical part of our Gaussian surface is zero because $\vec{E} \cdot \hat{n} = 0$ everywhere. The flux through each flat end of the surface is $+EA$ because $\vec{E} \cdot \hat{n} = E_\perp = E$ everywhere, so the total flux through both ends—and hence the total flux Φ_E through the Gaussian surface—is $+2EA$. The total enclosed charge is $Q_{\text{encl}} = \sigma A$, and so from Gauss's law,

$$2EA = \frac{\sigma A}{\epsilon_0} \quad \text{and}$$

$$E = \frac{\sigma}{2\epsilon_0} \qquad \text{(field of an infinite sheet of charge)}$$

If σ is negative, $\vec{E}$ is directed *toward* the sheet, the flux through the Gaussian surface in Fig. 22.20 is negative, and σ in the expression $E = \sigma/2\epsilon_0$ denotes the magnitude (absolute value) of the charge density.

Figure 22.20 A cylindrical Gaussian surface is used to find the field of an infinite plane sheet of charge.



EVALUATE We got the same result for the field of an infinite sheet of charge in Example 21.11 (Section 21.5). That calculation was much more complex and involved a fairly challenging integral. Thanks to the favorable symmetry, Gauss's law makes it much easier to solve this problem.

KEYCONCEPT To apply Gauss's law to a charge distribution with planar symmetry, such as an infinite sheet of charge, use a cylindrical Gaussian surface. At all points on the flat ends of the surface, the electric field has the same magnitude and is normal to the surface. On the curved sides of the cylindrical surface, the electric field is parallel to the surface, so the flux through the sides is zero.

EXAMPLE 22.8 Field between oppositely charged parallel conducting plates

WITH VARIATION PROBLEMS

Two large plane parallel conducting plates are given charges of equal magnitude and opposite sign; the surface charge densities are $+\sigma$ and $-\sigma$. Find the electric field in the region between the plates.

IDENTIFY and SET UP **Figure 22.21a** shows the field. Because opposite charges attract, most of the charge accumulates at the opposing faces of the plates. A small amount of charge resides on the *outer* surfaces of the plates, and there is some spreading or "fringing" of the field at the edges. But if the plates are very large in comparison to the distance between them, the amount of charge on the outer surfaces is negligibly small, and the fringing can be ignored except near the edges. In this case we can assume that the field is uniform in the interior region between the plates, as in Fig. 22.21b, and that the charges are distributed uniformly over the opposing surfaces. To exploit this symmetry, we can use the shaded Gaussian surfaces S_1, S_2, S_3, and S_4. These surfaces are cylinders with flat ends of area A; one end of each surface lies *within* a plate.

EXECUTE The left-hand end of surface S_1 is within the positive plate 1. Since the field is zero within the volume of any solid conductor under electrostatic conditions, there is no electric flux through this end. The electric field between the plates is perpendicular to the right-hand end, so on that end, $E_\perp$ is equal to E and the flux is EA; this is positive, since $\vec{E}$ is directed out of the Gaussian surface. There is no flux through the side walls of the cylinder, since these walls are parallel to $\vec{E}$. So the total flux integral in Gauss's law is EA. The net charge enclosed by the cylinder is σA, so Eq. (22.8) yields $EA = \sigma A/\epsilon_0$; we then have

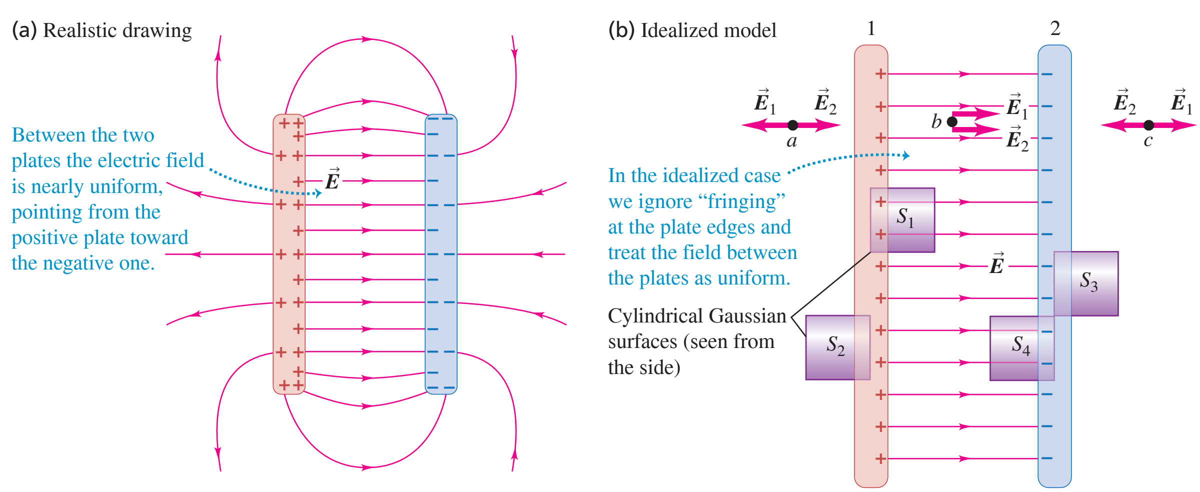
$$E = \frac{\sigma}{\epsilon_0} \quad \text{(field between oppositely charged conducting plates)}$$

The field is uniform and perpendicular to the plates, and its magnitude is independent of the distance from either plate. The Gaussian surface S_4 yields the same result. Surfaces S_2 and S_3 yield $E = 0$ to the left of plate 1 and to the right of plate 2, respectively. We leave these calculations to you.

EVALUATE We obtained the same results in Example 21.12 by using the principle of superposition of electric fields. The fields due to the two sheets of charge (one on each plate) are $\vec{E}_1$ and $\vec{E}_2$; from Example 22.7, both of these have magnitude $\sigma/2\epsilon_0$. The total electric field at any point is the vector sum $\vec{E} = \vec{E}_1 + \vec{E}_2$. At points a and c in Fig. 22.21b, $\vec{E}_1$ and $\vec{E}_2$ point in opposite directions, and their sum is zero. At point b, $\vec{E}_1$ and $\vec{E}_2$ are in the same direction; their sum has magnitude $E = \sigma/\epsilon_0$, just as we found by using Gauss's law.

KEYCONCEPT The electric field between two large, parallel, oppositely charged conducting plates is uniform in the region between the plates. The field is zero in the region outside the plates as well as in the interiors of the plates. The charge on either plate lies on the surface that faces the other plate.

Figure 22.21 Electric field between oppositely charged parallel plates.



EXAMPLE 22.9 Field of a uniformly charged sphere

WITH VARIATION PROBLEMS

Positive electric charge Q is distributed uniformly *throughout the volume* of an *insulating* sphere with radius R. Find the magnitude of the electric field at a point P a distance r from the center of the sphere.

IDENTIFY and SET UP As in Example 22.5, the system is spherically symmetric. Hence we can use the conclusions of that example about the direction and magnitude of $\vec{E}$. To make use of the spherical symmetry, we choose as our Gaussian surface a sphere with radius r, concentric with the charge distribution.

EXECUTE From symmetry, the direction of $\vec{E}$ is radial at every point on the Gaussian surface, so $E_\perp = E$ and the field magnitude E is the same at every point on the surface. Hence the total electric flux through the Gaussian surface is the product of E and the total area of the surface $A = 4\pi r^2$—that is, $\Phi_E = 4\pi r^2 E$.

The amount of charge enclosed within the Gaussian surface depends on r. To find E *inside* the sphere, we choose $r < R$. The volume charge density ρ is the charge Q divided by the volume of the entire charged sphere of radius R:

$$\rho = \frac{Q}{4\pi R^3/3}$$

The volume V_{encl} enclosed by the Gaussian surface is $\frac{4}{3}\pi r^3$, so the total charge Q_{encl} enclosed by that surface is

$$Q_{\text{encl}} = \rho V_{\text{encl}} = \left(\frac{Q}{4\pi R^3/3}\right)\left(\tfrac{4}{3}\pi r^3\right) = Q\frac{r^3}{R^3}$$

Then Gauss's law, Eq. (22.8), becomes

$$4\pi r^2 E = \frac{Q}{\epsilon_0}\frac{r^3}{R^3} \quad \text{and}$$

$$E = \frac{1}{4\pi\epsilon_0}\frac{Qr}{R^3} \quad \text{(field inside a uniformly charged sphere)}$$

The field magnitude is proportional to the distance r of the field point from the center of the sphere (see the graph of E versus r in **Fig. 22.22**).

To find E *outside* the sphere, we take $r > R$. This surface encloses the entire charged sphere, so $Q_{encl} = Q$, and Gauss's law gives

$$4\pi r^2 E = \frac{Q}{\epsilon_0} \quad \text{and}$$

$$E = \frac{1}{4\pi\epsilon_0}\frac{Q}{r^2} \quad \text{(field outside a uniformly charged sphere)}$$

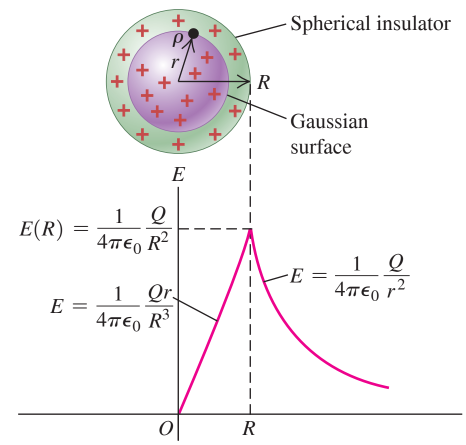
The field outside *any* spherically symmetric charged body varies as $1/r^2$, as though the entire charge were concentrated at the center. This is graphed in Fig. 22.22.

If the charge is *negative*, $\vec{E}$ is radially *inward* and in the expressions for E we interpret Q as the absolute value of the charge.

EVALUATE Notice that if we set $r = R$ in either expression for E, we get the same result $E = Q/4\pi\epsilon_0 R^2$ for the magnitude of the field at the surface of the sphere. This is because the magnitude E is a *continuous* function of r. By contrast, for the charged conducting sphere of Example 22.5 the electric-field magnitude is *discontinuous* at $r = R$ (it jumps from $E = 0$ just inside the sphere to $E = Q/4\pi\epsilon_0 R^2$ just outside the sphere). In general, the electric field $\vec{E}$ is discontinuous in magnitude, direction, or both wherever there is a *sheet* of charge, such as at the surface of a charged conducting sphere (Example 22.5), at the surface of an infinite charged sheet (Example 22.7), or at the surface of a charged conducting plate (Example 22.8).

The approach used here can be applied to *any* spherically symmetric distribution of charge, even if it is not radially uniform, as it was here.

Figure **22.22** The magnitude of the electric field of a uniformly charged insulating sphere. Compare this with the field for a conducting sphere (see Fig. 22.18).



Such charge distributions occur within many atoms and atomic nuclei, so Gauss's law is useful in atomic and nuclear physics.

KEYCONCEPT For any spherically symmetric distribution of charge, the field at a distance r from the center of the distribution is the same as if all the charge interior to r were concentrated in a point at the center. The charge at distances greater than r from the center has no effect on the field at r.

EXAMPLE 22.10 Charge on a hollow sphere

WITH VARIATION PROBLEMS

A thin-walled, hollow sphere of radius 0.250 m has an unknown charge distributed uniformly over its surface. At a distance of 0.300 m from the center of the sphere, the electric field points radially inward and has magnitude 1.80×10^2 N/C. How much charge is on the sphere?

IDENTIFY and SET UP The charge distribution is spherically symmetric. As in Examples 22.5 and 22.9, it follows that the electric field is radial everywhere and its magnitude is a function of only the radial distance r from the center of the sphere. We use a spherical Gaussian surface that is concentric with the charge distribution and has radius $r = 0.300$ m. Our target variable is $Q_{encl} = q$.

EXECUTE The charge distribution is the same as if the charge were on the surface of a 0.250-m-radius conducting sphere. Hence we can borrow the results of Example 22.5. We note that the electric field here is directed toward the sphere, so that q must be *negative*. Furthermore, the electric field is directed into the Gaussian surface, so that $E_\perp = -E$ and $\Phi_E = \oint E_\perp\, dA = -E(4\pi r^2)$.

By Gauss's law, the flux is equal to the charge q on the sphere (all of which is enclosed by the Gaussian surface) divided by ϵ_0. Solving for q, we find

$$\begin{aligned} q = -E(4\pi\epsilon_0 r^2) &= -(1.80 \times 10^2\ \text{N/C})(4\pi) \\ &\quad \times (8.854 \times 10^{-12}\ \text{C}^2/\text{N}\cdot\text{m}^2)(0.300\ \text{m})^2 \\ &= -1.80 \times 10^{-9}\ \text{C} = -1.80\ \text{nC} \end{aligned}$$

EVALUATE To determine the charge, we had to know the electric field at *all* points on the Gaussian surface so that we could calculate the flux integral. This was possible here because the charge distribution is highly symmetric. If the charge distribution is irregular or lacks symmetry, Gauss's law is not very useful for calculating the charge distribution from the field, or vice versa.

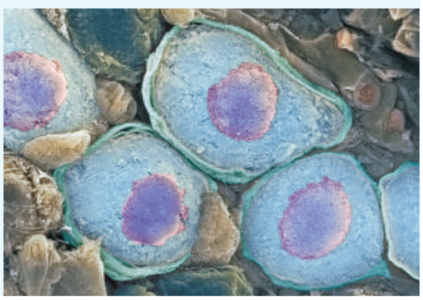
KEYCONCEPT The electric field outside a spherical distribution of charge is the same as if all of the charge in the distribution were concentrated at its center.

TEST YOUR UNDERSTANDING OF SECTION 22.4 You place a known amount of charge Q on the irregularly shaped conductor shown in Fig. 22.17. If you know the size and shape of the conductor, can you use Gauss's law to calculate the electric field at an arbitrary position outside the conductor?

ANSWER

no | You might be tempted to draw a Gaussian surface that is an enlarged version of the conductor, with the same shape and placed so that it completely encloses the conductor. While you know the flux through this Gaussian surface (by Gauss's law, it's $\Phi_E = Q/\epsilon_0$), the direction of the electric field need not be perpendicular to the surface and the magnitude of the field need not be the same at all points on the surface. It's not possible to do the flux integral $\oint E_\perp\, dA$, and we can't calculate the electric field. Gauss's law is useful for calculating the electric field only when the charge distribution is *highly* symmetric.

BIO APPLICATION Charge Distribution Inside a Nerve Cell The interior of a human nerve cell contains both positive potassium ions (K^+) and negatively charged protein molecules (Pr^-). Potassium ions can flow out of the cell through the cell membrane, but the much larger protein molecules cannot. The result is that the interior of the cell has a net negative charge. (The fluid outside the cell has a positive charge that balances this.) The fluid within the cell is a good conductor, so the Pr^- molecules distribute themselves on the outer surface of the fluid—that is, on the inner surface of the cell membrane, which is an insulator. This is true no matter what the shape of the cell.

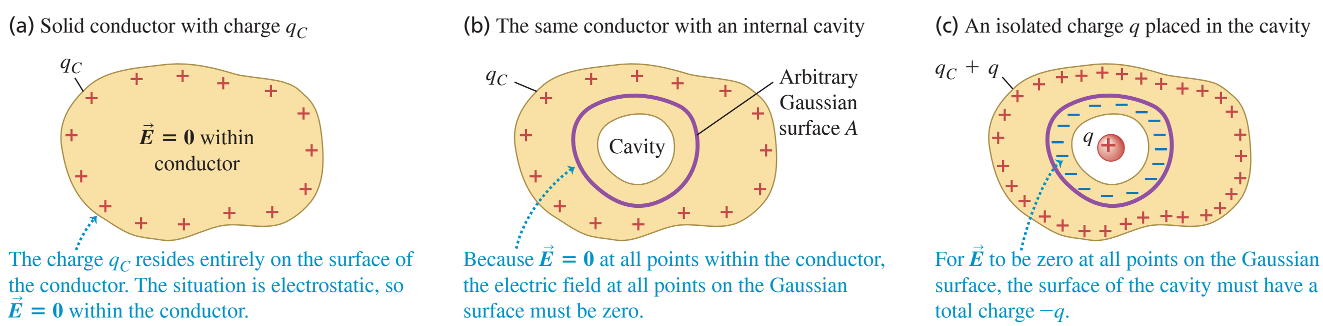


22.5 CHARGES ON CONDUCTORS

We have learned that in an electrostatic situation (in which there is no net motion of charge) the electric field at every point within a conductor is zero and any excess charge on a solid conductor is located entirely on its surface (**Fig. 22.23a**). But what if there is a *cavity* inside the conductor (Fig. 22.23b)? If there is no charge within the cavity, we can use a Gaussian surface such as A (which lies completely within the material of the conductor) to show that the *net* charge on the *surface of the cavity* must be zero, because $\vec{E} = \mathbf{0}$ everywhere on the Gaussian surface. In fact, we can prove in this situation that there can't be any charge *anywhere* on the cavity surface. We'll postpone detailed proof of this statement until Chapter 23.

Suppose we place a small body with a charge q inside a cavity within a conductor (Fig. 22.23c). The conductor is uncharged and is insulated from the charge q. Again $\vec{E} = \mathbf{0}$ everywhere on surface A, so according to Gauss's law the *total* charge inside this surface must be zero. Therefore there must be a charge $-q$ distributed on the surface of the cavity, drawn there by the charge q inside the cavity. The *total* charge on the conductor must remain zero, so a charge $+q$ must appear either on its outer surface or inside the material. But we showed that in an electrostatic situation there can't be any excess charge within the material of a conductor. So we conclude that charge $+q$ must appear on the outer surface. By the same reasoning, if the conductor originally had a charge q_C, then the total charge on the outer surface must be $q_C + q$ after charge q is inserted into the cavity.

Figure **22.23** Finding the electric field within a charged conductor.



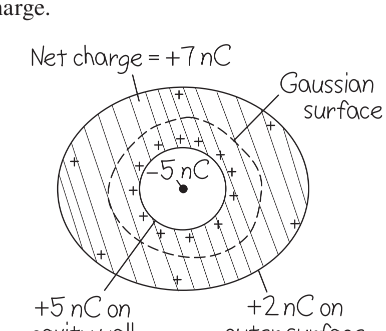
CONCEPTUAL EXAMPLE 22.11 A conductor with a cavity

WITH VARIATION PROBLEMS

A conductor with a cavity carries a total charge of +7 nC. Within the cavity, insulated from the conductor, is a point charge of −5 nC. How much charge is on each surface (inner and outer) of the conductor?

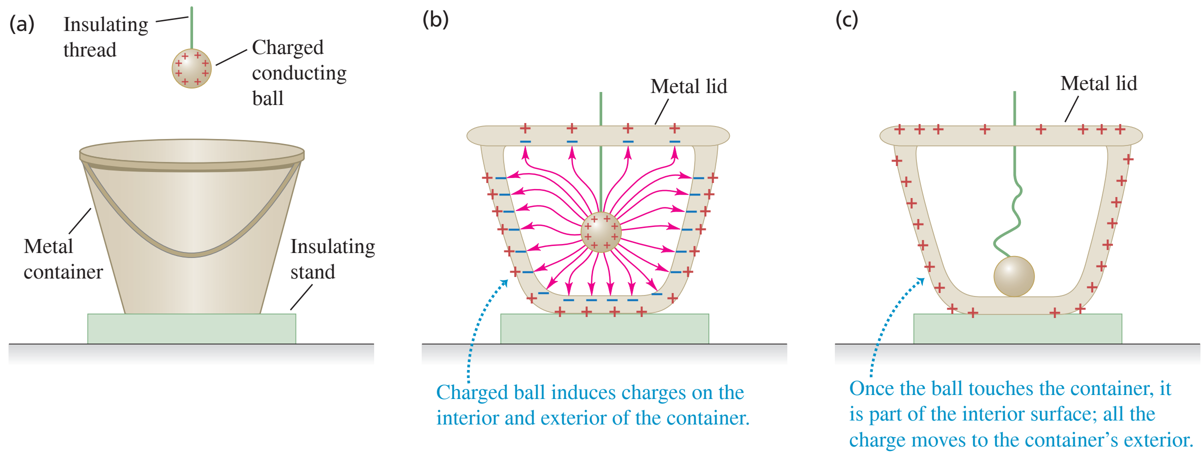
SOLUTION **Figure 22.24** shows the situation. If the charge in the cavity is $q = -5$ nC, the charge on the inner cavity surface must be $-q = -(-5\text{ nC}) = +5$ nC. The conductor carries a *total* charge of +7 nC, none of which is in the interior of the material. If +5 nC is on the inner surface of the cavity, then there must be $(+7\text{ nC}) - (+5\text{ nC}) = +2$ nC on the outer surface of the conductor.

Figure **22.24** Our sketch for this problem. There is zero electric field inside the bulk conductor and hence zero flux through the Gaussian surface shown, so the charge on the cavity wall must be the opposite of the point charge.



KEYCONCEPT If there is excess charge at rest on a conductor, all of that charge must lie on the surface of the conductor and the electric field inside the solid conductor must be zero. If there is a cavity inside the conductor, the net charge on the cavity walls equals the amount of charge enclosed by the cavity.

Figure **22.25** (a) A charged conducting ball suspended by an insulating thread outside a conducting container on an insulating stand. (b) The ball is lowered into the container, and the lid is put on. (c) The ball is touched to the inner surface of the container.



Testing Gauss's Law Experimentally

Figure 22.25 shows a historic experiment designed to test Gauss's law. We mount a conducting container on an insulating stand. The container is initially uncharged. Then we hang a charged metal ball from an insulating thread (Fig. 22.25a), lower it into the container, and put the lid on (Fig. 22.25b). Charges are induced on the walls of the container, as shown. But now we let the ball *touch* the inner wall (Fig. 22.25c). The surface of the ball becomes part of the cavity surface. The situation is now the same as Fig. 22.23b; if Gauss's law is correct, the net charge on the cavity surface must be zero. Thus the ball must lose all its charge. Finally, we pull the ball out; we find that it has indeed lost all its charge.

This experiment was performed in the 19th century by the English scientist Michael Faraday, using a metal icepail with a lid, and it is called **Faraday's icepail experiment.** The result confirms the validity of Gauss's law and therefore of Coulomb's law. Faraday's result was significant because Coulomb's experimental method, using a torsion balance and dividing of charges, was not very precise; it is very difficult to confirm the $1/r^2$ dependence of the electrostatic force by direct force measurements. By contrast, experiments like Faraday's test the validity of Gauss's law, and therefore of Coulomb's law, with much greater precision. Modern versions of this experiment have shown that the exponent 2 in the $1/r^2$ of Coulomb's law does not differ from precisely 2 by more than 10^{-16}. So there is no reason to believe it is anything other than exactly 2.

The same principle behind Faraday's icepail experiment is used in a *Van de Graaff electrostatic generator* (**Fig. 22.26**). A charged belt continuously produces a buildup of charge on the inside of a conducting shell. By Gauss's law, there can never be any charge on the inner surface of this shell, so the charge is immediately carried away to the outside surface of the shell. As a result, the charge on the shell and the electric field around it can become very large very rapidly. The Van de Graaff generator is used as an accelerator of charged particles and for physics demonstrations.

This principle also forms the basis for *electrostatic shielding*. Suppose we have a very sensitive electronic instrument that we want to protect from stray electric fields that might cause erroneous measurements. We surround the instrument with a conducting box, or we line the walls, floor, and ceiling of the room with a conducting material such as sheet copper. The external electric field redistributes the free electrons in the conductor, leaving a net positive charge on the outer surface in some regions and a net negative charge in others (**Fig. 22.27**, next page). This charge distribution causes an additional electric field such that the *total* field at every point inside the box is zero, as Gauss's law says it must be. The charge distribution on the box also alters the shapes of the field lines near the box, as the figure shows. Such a setup is often called a *Faraday cage*. The same

Figure **22.26** Cutaway view of the essential parts of a Van de Graaff electrostatic generator. The electron sink at the bottom draws electrons from the belt, giving it a positive charge; at the top the belt attracts electrons away from the conducting shell, giving the shell a positive charge.

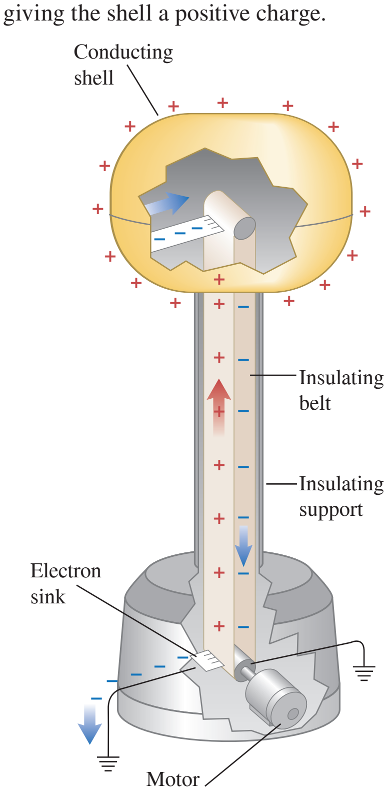
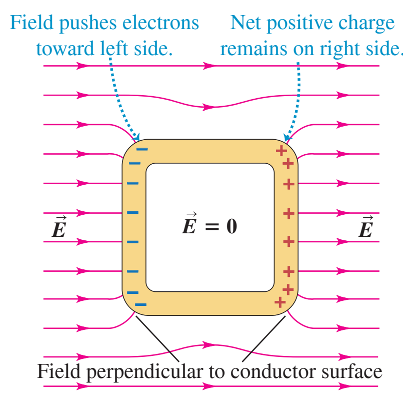


Figure **22.27** (a) A conducting box (a Faraday cage) immersed in a uniform electric field. The field of the induced charges on the box combines with the uniform field to give zero total field inside the box. (b) This person is inside a Faraday cage, and so is protected from the powerful electric discharge.

(a)



(b)

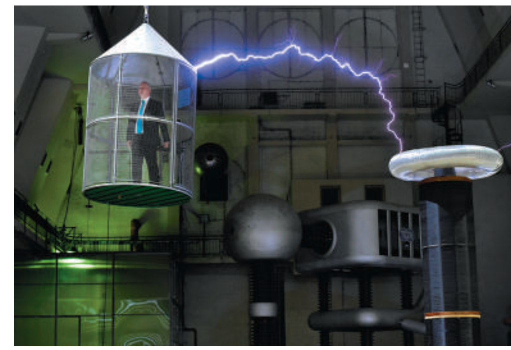
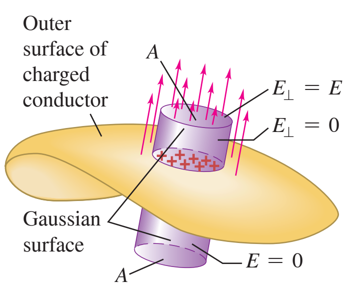


Figure **22.28** The field just outside a charged conductor is perpendicular to the surface, and its perpendicular component $E_\perp$ is equal to σ/ϵ_0.



physics tells you that one of the safest places to be in a lightning storm is inside a car; if the car is struck by lightning, the charge tends to remain on the metal skin of the vehicle, and little or no electric field is produced inside the passenger compartment. Keep in mind this basic principle, which is a consequence of Gauss's law: *Charges outside a conductor have no effect on the interior of that conductor, even if the conductor has a cavity inside.*

Field at the Surface of a Conductor

Finally, we note that there is a direct relationship between the $\vec{E}$ field at a point just outside any conductor and the surface charge density σ at that point. In general, σ varies from point to point on the surface. We'll show in Chapter 23 that at any such point, the direction of $\vec{E}$ is always *perpendicular* to the surface. (You can see this effect in Fig. 22.27a.)

To find a relationship between σ at any point on the surface and the perpendicular component of the electric field at that point, we construct a Gaussian surface in the form of a small cylinder (**Fig. 22.28**). One end face, with area A, lies within the conductor and the other lies just outside. The electric field is zero at all points within the conductor. Outside the conductor the component of $\vec{E}$ perpendicular to the side walls of the cylinder is zero, and over the end face the perpendicular component is equal to $E_\perp$. (If σ is positive, the electric field points out of the conductor and $E_\perp$ is positive; if σ is negative, the field points inward and $E_\perp$ is negative.) Hence the total flux through the surface is $E_\perp A$. The charge enclosed within the Gaussian surface is σA, so from Gauss's law, $E_\perp A = (\sigma A)/\epsilon_0$ and

Electric field at surface of a conductor, $\vec{E}$ perpendicular to surface

$$E_\perp = \frac{\sigma}{\epsilon_0} \tag{22.10}$$

Surface charge density (σ); Electric constant (ϵ_0)

This agrees with our result for the field at the surface of a charged conducting plate (Example 22.8). We can also verify this for a charged conducting sphere. In Example 22.5 (Section 22.4) we showed that the electric field just outside the surface of a conducting sphere with radius R and total charge q is

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{R^2}$$

The surface charge density is uniform and equal to q divided by the surface area of the sphere:

$$\sigma = \frac{q}{4\pi R^2}$$

Comparing these two expressions, we see that $E = \sigma/\epsilon_0$, which verifies Eq. (22.10).

APPLICATION Why Lightning Bolts Are Vertical Our planet is a good conductor, and its surface has a negative charge. Hence, the electric field in the atmosphere above the surface points generally downward, toward the negative charge and perpendicular to the surface (see Example 22.12). The negative charge is balanced by positive charges in the atmosphere. In a lightning storm, the vertical electric field becomes great enough to cause charges to flow vertically through the air. The air is excited and ionized by the passage of charge through it, producing a visible lightning bolt.

We showed in Example 22.8 that the field magnitude between two infinite flat oppositely charged conducting plates also equals σ/ϵ_0. In this case the field magnitude E is the same at *all* distances from the plates, but in all other cases E decreases with increasing distance from the surface.

EXAMPLE 22.12 Electric field of the earth

WITH VARIATION PROBLEMS

The earth (a conductor) has a net electric charge. The resulting electric field near the surface has an average value of about 150 N/C, directed toward the center of the earth. (a) What is the corresponding surface charge density? (b) What is the *total* surface charge of the earth?

IDENTIFY and SET UP We are given the electric-field magnitude at the surface of the conducting earth. We can calculate the surface charge density σ from Eq. (22.10). The total charge Q on the earth's surface is then the product of σ and the earth's surface area.

EXECUTE (a) The direction of the field means that σ is negative (corresponding to $\vec{E}$ being directed *into* the surface, so $E_\perp$ is negative). From Eq. (22.10),

$$\begin{aligned}\sigma &= \epsilon_0 E_\perp = (8.85 \times 10^{-12}\ \mathrm{C^2/N \cdot m^2})(-150\ \mathrm{N/C})\\ &= -1.33 \times 10^{-9}\ \mathrm{C/m^2} = -1.33\ \mathrm{nC/m^2}\end{aligned}$$

(b) The earth's surface area is $4\pi R_\mathrm{E}^2$, where $R_\mathrm{E} = 6.38 \times 10^6$ m is the radius of the earth (see Appendix F). The total charge Q is the product $4\pi R_\mathrm{E}^2 \sigma$, or

$$\begin{aligned}Q &= 4\pi(6.38 \times 10^6\ \mathrm{m})^2(-1.33 \times 10^{-9}\ \mathrm{C/m^2})\\ &= -6.8 \times 10^5\ \mathrm{C} = -680\ \mathrm{kC}\end{aligned}$$

EVALUATE You can check our result in part (b) by using the result of Example 22.5. Solving for Q, we find

$$\begin{aligned}Q &= 4\pi\epsilon_0 R^2 E_\perp\\ &= \frac{1}{9.0 \times 10^9\ \mathrm{N \cdot m^2/C^2}}(6.38 \times 10^6\ \mathrm{m})^2(-150\ \mathrm{N/C})\\ &= -6.8 \times 10^5\ \mathrm{C}\end{aligned}$$

One electron has a charge of -1.60×10^{-19} C. Hence this much excess negative electric charge corresponds to there being $(-6.8 \times 10^5\ \mathrm{C})/(-1.60 \times 10^{-19}\ \mathrm{C}) = 4.2 \times 10^{24}$ excess electrons on the earth, or about 7 moles of excess electrons. This is compensated by an equal *deficiency* of electrons in the earth's upper atmosphere, so the combination of the earth and its atmosphere is electrically neutral.

KEYCONCEPT The electric field $\vec{E}$ at the surface of a charged conductor is normal to the surface. The normal component of $\vec{E}$ at any point on the surface is equal to σ/ϵ_0, where σ is the surface charge density (charge per unit area) at that point. The field points out of the surface if σ is positive and into the surface if σ is negative.

TEST YOUR UNDERSTANDING OF SECTION 22.5 A hollow conducting sphere has no net charge. There is a positive point charge q at the center of the spherical cavity within the sphere. You connect a conducting wire from the outside of the sphere to ground. Will you measure an electric field outside the sphere?

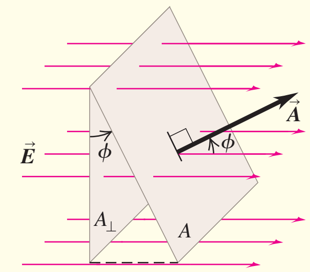
ANSWER

no Before you connect the wire to the sphere, the presence of the point charge will induce a charge $-q$ on the inner surface of the hollow sphere and a charge q on the outer surface (the net charge on the sphere is zero). There will be an electric field outside the sphere due to the charge on the outer surface. Once you touch the conducting wire to the sphere, however, electrons will flow from ground to the outer surface of the sphere to neutralize the charge there (see Fig. 21.7c). As a result the sphere will have no charge on its outer surface and no electric field outside.

CHAPTER 22 SUMMARY

Electric flux: Electric flux is a measure of the "flow" of electric field through a surface. It is equal to the product of an area element and the perpendicular component of $\vec{E}$, integrated over a surface. (See Examples 22.1–22.3.)

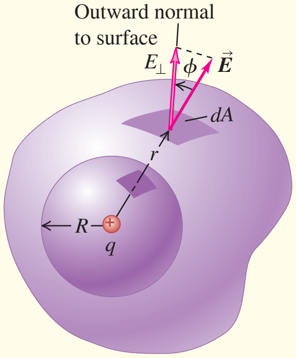
$$\Phi_E = \int E\cos\phi\, dA = \int E_\perp\, dA = \int \vec{E}\cdot d\vec{A} \qquad (22.5)$$



Gauss's law: Gauss's law states that the total electric flux through a closed surface, which can be written as the surface integral of the component of $\vec{E}$ normal to the surface, equals a constant times the total charge Q_{encl} enclosed by the surface. Gauss's law is logically equivalent to Coulomb's law, but its use greatly simplifies problems with a high degree of symmetry. (See Examples 22.4–22.10.)

When excess charge is placed on a solid conductor and is at rest, it resides entirely on the surface, and $\vec{E} = \mathbf{0}$ everywhere in the material of the conductor. (See Examples 22.11 and 22.12.)

$$\Phi_E = \oint E\cos\phi\, dA = \oint E_\perp\, dA = \oint \vec{E}\cdot d\vec{A} = \frac{Q_{\text{encl}}}{\epsilon_0} \qquad (22.8), (22.9)$$



Electric field of various symmetric charge distributions: The following table lists electric fields caused by several symmetric charge distributions. In the table, q, Q, λ, and σ refer to the *magnitudes* of the quantities.

Charge Distribution	Point in Electric Field	Electric-Field Magnitude
Single point charge q	Distance r from q	$E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}$
Charge q on surface of conducting sphere with radius R	Outside sphere, $r > R$	$E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}$
	Inside sphere, $r < R$	$E = 0$
Infinite wire, charge per unit length λ	Distance r from wire	$E = \frac{1}{2\pi\epsilon_0}\frac{\lambda}{r}$
Infinite conducting cylinder with radius R, charge per unit length λ	Outside cylinder, $r > R$	$E = \frac{1}{2\pi\epsilon_0}\frac{\lambda}{r}$
	Inside cylinder, $r < R$	$E = 0$
Solid insulating sphere with radius R, charge Q distributed uniformly throughout volume	Outside sphere, $r > R$	$E = \frac{1}{4\pi\epsilon_0}\frac{Q}{r^2}$
	Inside sphere, $r < R$	$E = \frac{1}{4\pi\epsilon_0}\frac{Qr}{R^3}$
Infinite sheet of charge with uniform charge per unit area σ	Any point	$E = \frac{\sigma}{2\epsilon_0}$
Two oppositely charged conducting plates with surface charge densities $+\sigma$ and $-\sigma$	Any point between plates	$E = \frac{\sigma}{\epsilon_0}$
Charged conductor	Just outside the conductor	$E = \frac{\sigma}{\epsilon_0}$

GUIDED PRACTICE

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 22.2 and 22.3 (Section 22.2) and EXAMPLE 22.4 (Section 22.3) before attempting these problems.

VP22.4.1 An imaginary cubical surface of side L has its edges parallel to the x-, y- and z-axes, one corner at the point $x = 0$, $y = 0$, $z = 0$ and the opposite corner at the point $x = L$, $y = L$, $z = L$. The cube is in a region of uniform electric field $\vec{E} = E_1\hat{\imath} + E_2\hat{\jmath}$, where E_1 and E_2 are positive constants. Calculate the electric flux through (a) the cube face in the plane $x = 0$, (b) the cube face in the plane $x = L$, (c) the cube face in the plane $y = 0$, (d) the cube face in the plane $y = L$, (e) the cube face in the plane $z = 0$, (f) the cube face in the plane $z = L$, and (g) the entire cubical surface. For each face the normal points out of the cube.

VP22.4.2 A point charge q is at the point $x = 0$, $y = 0$, $z = 0$. An imaginary hemispherical surface is made by starting with a spherical surface of radius R centered on the point $x = 0$, $y = 0$, $z = 0$ and cutting off the half in the region $z < 0$. The normal to this surface points out of the hemisphere, away from its center. Calculate the electric flux through the hemisphere if (a) $q = 8.00$ nC and $R = 0.100$ m; (b) $q = -4.00$ nC and $R = 0.100$ m; (c) $q = -4.00$ nC and $R = 0.200$ m.

VP22.4.3 An imaginary cubical surface with sides of length 5.00 cm has a point charge $q = +6.00$ nC at its center. Calculate the electric flux (a) through the entire closed cubical surface and (b) through one of the six faces of the cube.

VP22.4.4 An imaginary spherical surface of radius 5.00 cm is centered on the point $x = 0$, $y = 0$, $z = 0$. Calculate the net electric flux through the surface if the following charges are present: (a) a point charge $q_1 = +3.00$ nC at $x = 1.00$ cm, $y = 1.00$ cm, $z = 0$; (b) the charge in part (a) plus a point charge $q_2 = -8.00$ nC at $x = 2.00$ cm, $y = 0$, $z = -4.00$ cm; (c) the charges in part (b) plus a point charge $q_3 = +2.00$ nC at $x = 4.00$ cm, $y = -2.00$ cm, $z = 3.00$ cm.

Be sure to review EXAMPLES 22.5, 22.6, 22.7, 22.8, 22.9, and 22.10 (Section 22.4) before attempting these problems.

VP22.10.1 A sphere of radius 5.00 cm carries charge +3.00 nC. Calculate the electric-field magnitude at a distance 4.00 cm from the center of the sphere and at a distance 6.00 cm from the center of the sphere if the sphere is (a) a solid insulator with the charge spread uniformly throughout its volume and (b) a solid conductor.

VP22.10.2 You place a point charge $q = -4.00$ nC a distance of 9.00 cm from an infinitely long, thin wire that has linear charge density 3.00×10^{-9} C/m. What is the magnitude of the electric force that the wire exerts on the point charge? Is the force attractive or repulsive?

VP22.10.3 A point charge $q = +1.80$ nC is 8.00 cm from a thin, flat, infinite sheet that has a uniform surface charge density. If the total electric field is zero at a point halfway between the point charge and the sheet, what is the surface charge density of the sheet? Is this positive or negative?

VP22.10.4 Two large parallel conducting plates carry charges of equal magnitude and opposite charge. When you place a point charge $q = +3.60$ nC between the plates, the force on the point charge is 22.0 μN. What is the magnitude of the surface charge density on either plate?

Be sure to review EXAMPLES 22.11 and 22.12 (Section 22.5) before attempting these problems.

VP22.12.1 A solid block of copper, which is a good conductor, has a cavity in its interior. Within the cavity, insulated from the conductor, is a point charge of +3.00 nC. The block of copper itself has an excess charge of −8.00 nC. How much charge is on (a) the outer surface of the block and (b) the surface of the cavity?

VP22.12.2 A solid sphere of silver, which is a good conductor, has a spherical cavity at its center. There is a point charge at the center of the cavity. The silver sphere has a charge of +9.00 nC on its outer surface and a charge of −2.00 nC on the surface of the cavity. (a) What is the value of the point charge? (b) If the point charge moved to a different position within the cavity (not at the center), would this affect the total charge on the surface of the cavity or the total charge on the outer surface of the sphere?

VP22.12.3 In an electricity demonstration at the Deutsches Museum in Munich, Germany, a person sits inside a metal sphere of radius 0.90 m. Charge is applied to the sphere until the magnitude of the electric field on the outside surface of the sphere is 3.0×10^5 N/C. When the demonstration is over, the person climbs out unharmed. During the demonstration, what are the magnitudes of (a) the electric field inside the sphere, (b) the total charge on the sphere, and (c) the surface charge density on the outside of the sphere?

VP22.12.4 A spherical conductor of radius 0.330 m has a spherical cavity of radius 0.120 m at its center. The conductor carries a total charge of −6.00 nC; in addition, at the center of the spherical cavity is a point charge of +4.00 nC. Find (a) the total charge on the surface of the cavity, (b) the total charge on the outer surface of the conductor, (c) the magnitude of the electric field just inside the surface of the cavity, and (d) the magnitude of the electric field just outside the outer surface of the conductor.

BRIDGING PROBLEM Electric Field Inside a Hydrogen Atom

A hydrogen atom is made up of a proton of charge $+Q = 1.60 \times 10^{-19}$ C and an electron of charge $-Q = -1.60 \times 10^{-19}$ C. The proton may be regarded as a point charge at $r = 0$, the center of the atom. The motion of the electron causes its charge to be "smeared out" into a spherical distribution around the proton (**Fig. 22.29**), so that the electron is equivalent to a charge per unit volume of $\rho(r) = -(Q/\pi a_0^3)e^{-2r/a_0}$, where $a_0 = 5.29 \times 10^{-11}$ m is called the *Bohr radius.* (a) Find the total amount of the hydrogen atom's charge that is enclosed within a sphere with radius r centered on the proton. (b) Find the electric field (magnitude and direction) caused by the charge of the hydrogen atom as a function of r. (c) Make a graph as a function of r of the ratio of the electric-field magnitude E to the magnitude of the field due to the proton alone.

SOLUTION GUIDE

IDENTIFY and SET UP

1. The charge distribution in this problem is spherically symmetric, as in Example 22.9, so you can solve it with Gauss's law.
2. The charge within a sphere of radius r includes the proton charge $+Q$ plus the portion of the electron charge distribution

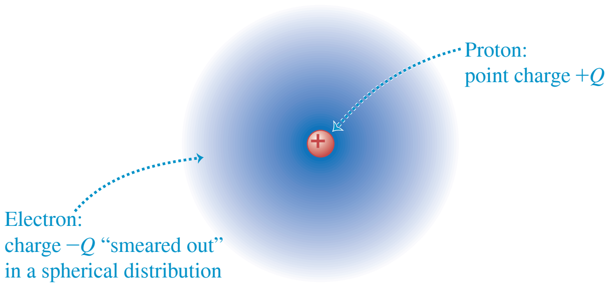


Figure 22.29 The charge distribution in a hydrogen atom.

that lies within the sphere. The difference from Example 22.9 is that the electron charge distribution is *not* uniform, so the charge enclosed within a sphere of radius r is *not* simply the charge density multiplied by the volume $4\pi r^3/3$ of the sphere. Instead, you'll have to do an integral.

3. Consider a thin spherical shell centered on the proton, with radius r' and infinitesimal thickness dr'. Since the shell is so thin, every point within the shell is at essentially the same radius from the proton. Hence the amount of electron charge within this shell *is* equal to the electron charge density $\rho(r')$ at this radius multiplied by the volume dV of the shell. What is dV in terms of r'?
4. The total electron charge within a radius r equals the integral of $\rho(r')dV$ from $r' = 0$ to $r' = r$. Set up this integral (but don't solve it yet), and use it to write an expression for the total charge (including the proton) within a sphere of radius r.

EXECUTE

5. Integrate your expression from step 4 to find the charge within radius r. (*Hint:* Integrate by substitution: Change the integration variable from r' to $x = 2r'/a_0$. You can use integration by parts to calculate the integral $\int x^2 e^{-x}\,dx$, or you can look it up in a table of integrals or on the Web.)
6. Use Gauss's law and your results from step 5 to find the electric field at a distance r from the proton.
7. Find the ratio referred to in part (c) and graph it versus r. (You'll actually find it simplest to graph this function versus the quantity r/a_0.)

EVALUATE

8. How do your results for the enclosed charge and the electric-field magnitude behave in the limit $r \to 0$? In the limit $r \to \infty$? Explain your results.

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q22.1 A rubber balloon has a single point charge in its interior. Does the electric flux through the balloon depend on whether or not it is fully inflated? Explain your reasoning.

Q22.2 Suppose that in Fig. 22.15 both charges were positive. What would be the fluxes through each of the four surfaces in the example?

Q22.3 In Fig. 22.15, suppose a third point charge were placed outside the purple Gaussian surface C. Would this affect the electric flux through any of the surfaces A, B, C, or D in the figure? Why or why not?

Q22.4 A certain region of space bounded by an imaginary closed surface contains no charge. Is the electric field always zero everywhere on the surface? If not, under what circumstances is it zero on the surface?

Q22.5 A spherical Gaussian surface encloses a point charge q. If the point charge is moved from the center of the sphere to a point away from the center, does the electric field at a point on the surface change? Does the total flux through the Gaussian surface change? Explain.

Q22.6 You find a sealed box on your doorstep. You suspect that the box contains several charged metal spheres packed in insulating material. How can you determine the total net charge inside the box without opening the box? Or isn't this possible?

Q22.7 A solid copper sphere has a net positive charge. The charge is distributed uniformly over the surface of the sphere, and the electric field inside the sphere is zero. Then a negative point charge outside the sphere is brought close to the surface of the sphere. Is all the net charge on the sphere still on its surface? If so, is this charge still distributed uniformly over the surface? If it is not uniform, how is it distributed? Is the electric field inside the sphere still zero? In each case justify your answers.

Q22.8 If the electric field of a point charge were proportional to $1/r^3$ instead of $1/r^2$, would Gauss's law still be valid? Explain your reasoning. (*Hint:* Consider a spherical Gaussian surface centered on a single point charge.)

Q22.9 In a conductor, one or more electrons from each atom are free to roam throughout the volume of the conductor. Does this contradict the statement that any excess charge on a solid conductor must reside on its surface? Why or why not?

Q22.10 You charge up the Van de Graaff generator shown in Fig. 22.26, and then bring an identical but uncharged hollow conducting sphere near it, without letting the two spheres touch. Sketch the distribution of charges on the second sphere. What is the net flux through the second sphere? What is the electric field inside the second sphere?

Q22.11 A lightning rod is a rounded copper rod mounted on top of a building and welded to a heavy copper cable running down into the ground. Lightning rods are used to protect houses and barns from lightning; the lightning current runs through the copper rather than through the building. Why? Why should the end of the rod be rounded?

Q22.12 A solid conductor has a cavity in its interior. Would the presence of a point charge inside the cavity affect the electric field outside the conductor? Why or why not? Would the presence of a point charge outside the conductor affect the electric field inside the cavity? Again, why or why not?

Q22.13 Explain this statement: "In a static situation, the electric field at the surface of a conductor can have no component parallel to the surface because this would violate the condition that the charges on the surface are at rest." Would this statement be valid for the electric field at the surface of an *insulator?* Explain your answer and the reason for any differences between the cases of a conductor and an insulator.

Q22.14 In a certain region of space, the electric field $\vec{E}$ is uniform. (a) Use Gauss's law to prove that this region of space must be electrically neutral; that is, the volume charge density ρ must be zero. (b) Is the converse true? That is, in a region of space where there is no charge, must $\vec{E}$ be uniform? Explain.

Q22.15 (a) In a certain region of space, the volume charge density ρ has a uniform positive value. Can $\vec{E}$ be uniform in this region? Explain. (b) Suppose that in this region of uniform positive ρ there is a "bubble" within which $\rho = 0$. Can $\vec{E}$ be uniform within this bubble? Explain.

Q22.16 A negative charge $-Q$ is placed inside the cavity of a hollow metal solid. The outside of the solid is grounded by connecting a conducting wire between it and the earth. Is any excess charge induced on the inner surface of the metal? Is there any excess charge on the outside surface of the metal? Why or why not? Would someone outside the solid measure an electric field due to the charge $-Q$? Is it reasonable to say that the grounded conductor has *shielded* the region outside the conductor from the effects of the charge $-Q$? In principle, could the same thing be done for gravity? Why or why not?

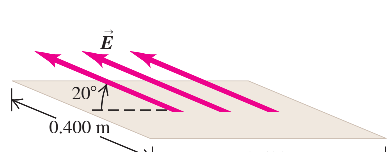
EXERCISES

Section 22.2 Calculating Electric Flux

22.1 • A flat sheet of paper of area $0.250\ \text{m}^2$ is oriented so that the normal to the sheet is at an angle of 60° to a uniform electric field of magnitude 14 N/C. (a) Find the magnitude of the electric flux through the sheet. (b) Does the answer to part (a) depend on the shape of the sheet? Why or why not? (c) For what angle ϕ between the normal to the sheet and the electric field is the magnitude of the flux through the sheet (i) largest and (ii) smallest? Explain your answers.

22.2 •• A flat sheet is in the shape of a rectangle with sides of lengths 0.400 m and 0.600 m. The sheet is immersed in a uniform electric field of magnitude 90.0 N/C that is directed at 20° from the plane of the sheet (**Fig. E22.2**). Find the magnitude of the electric flux through the sheet.

Figure **E22.2**



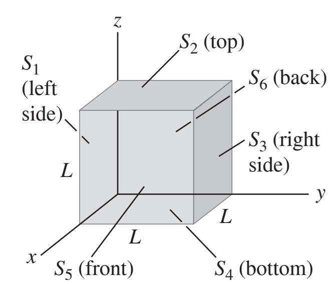
22.3 • You measure an electric field of 1.25×10^6 N/C at a distance of 0.150 m from a point charge. There is no other source of electric field in the region other than this point charge. (a) What is the electric flux through the surface of a sphere that has this charge at its center and that has radius 0.150 m? (b) What is the magnitude of this charge?

22.4 • It was shown in Example 21.10 (Section 21.5) that the electric field due to an infinite line of charge is perpendicular to the line and has magnitude $E = \lambda/2\pi\epsilon_0 r$. Consider an imaginary cylinder with radius $r = 0.250$ m and length $l = 0.400$ m that has an infinite line of positive charge running along its axis. The charge per unit length on the line is $\lambda = 3.00\ \mu\text{C/m}$. (a) What is the electric flux through the cylinder due to this infinite line of charge? (b) What is the flux through the cylinder if its radius is increased to $r = 0.500$ m? (c) What is the flux through the cylinder if its length is increased to $l = 0.800$ m?

22.5 • A uniform electric field makes an angle of 60.0° with a flat surface. The area of the surface is $6.66 \times 10^{-4}\ \text{m}^2$. The resulting electric flux through the surface is $4.44\ \text{N} \cdot \text{m}^2/\text{C}$. Calculate the magnitude of the electric field.

22.6 • The cube in **Fig. E22.6** has sides of length $L = 10.0$ cm. The electric field is uniform, has magnitude $E = 4.00 \times 10^3$ N/C, and is parallel to the xy-plane at an angle of 53.1° measured from the $+x$-axis toward the $+y$-axis. (a) What is the electric flux through each of the six cube faces S_1, S_2, S_3, S_4, S_5, and S_6? (b) What is the total electric flux through all faces of the cube?

Figure **E22.6**

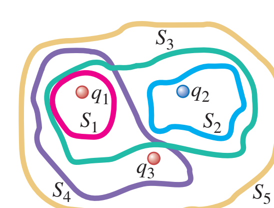


Section 22.3 Gauss's Law

22.7 • A charge of 87.6 pC is uniformly distributed on the surface of a thin sheet of insulating material that has a total area of $29.2\ \text{cm}^2$. A Gaussian surface encloses a portion of the sheet of charge. If the flux through the Gaussian surface is $5.00\ \text{N} \cdot \text{m}^2/\text{C}$, what area of the sheet is enclosed by the Gaussian surface?

22.8 • The three small spheres shown in **Fig. E22.8** carry charges $q_1 = 4.00$ nC, $q_2 = -7.80$ nC, and $q_3 = 2.40$ nC. Find the net electric flux through each of the following closed surfaces shown in cross section in the figure: (a) S_1; (b) S_2; (c) S_3; (d) S_4; (e) S_5. (f) Do your answers to parts (a)–(e) depend on how the charge is distributed over each small sphere? Why or why not?

Figure **E22.8**



Surface	What it encloses
S_1	q_1
S_2	q_2
S_3	q_1 and q_2
S_4	q_1 and q_3
S_5	q_1 and q_2 and q_3

22.9 •• A charged paint is spread in a very thin uniform layer over the surface of a plastic sphere of diameter 12.0 cm, giving it a charge of $-49.0\ \mu\text{C}$. Find the electric field (a) just inside the paint layer; (b) just outside the paint layer; (c) 5.00 cm outside the surface of the paint layer.

22.10 • A point charge $q_1 = 4.00$ nC is located on the x-axis at $x = 2.00$ m, and a second point charge $q_2 = -6.00$ nC is on the y-axis at $y = 1.00$ m. What is the total electric flux due to these two point charges through a spherical surface centered at the origin and with radius (a) 0.500 m, (b) 1.50 m, (c) 2.50 m?

22.11 • **BIO** As discussed in Section 22.5, human nerve cells have a net negative charge and the material in the interior of the cell is a good conductor. If a cell has a net charge of -8.65 pC, what are the magnitude and direction (inward or outward) of the net flux through the cell boundary?

22.12 • **Electric Fields in an Atom.** The nuclei of large atoms, such as uranium, with 92 protons, can be modeled as spherically symmetric spheres of charge. The radius of the uranium nucleus is approximately 7.4×10^{-15} m. (a) What is the electric field this nucleus produces just outside its surface? (b) What magnitude of electric field does it produce at the distance of the electrons, which is about 1.0×10^{-10} m? (c) The electrons can be modeled as forming a uniform shell of negative charge. What net electric field do they produce at the location of the nucleus?

Section 22.4 Applications of Gauss's Law and Section 22.5 Charges on Conductors

22.13 •• Two very long uniform lines of charge are parallel and are separated by 0.300 m. Each line of charge has charge per unit length $+5.20\ \mu$C/m. What magnitude of force does one line of charge exert on a 0.0500 m section of the other line of charge?

22.14 •• A solid metal sphere with radius 0.450 m carries a net charge of 0.250 nC. Find the magnitude of the electric field (a) at a point 0.100 m outside the surface of the sphere and (b) at a point inside the sphere, 0.100 m below the surface.

22.15 • (a) A conducting sphere has charge Q and radius R. If the electric field of the sphere at a distance $r = 2R$ from the center of the sphere is 1400 N/C, what is the electric field of the sphere at $r = 4R$? (b) A very long conducting cylinder of radius R has charge per unit length λ. Let r be the perpendicular distance from the axis of the cylinder. If the electric field of the cylinder at $r = 2R$ is 1400 N/C, what is the electric field at $r = 4R$? (c) A very large uniform sheet of charge has surface charge density σ. If the electric field of the sheet has a value of 1400 N/C at a perpendicular distance d from the sheet, what is the electric field of the sheet at a distance of $2d$ from the sheet?

22.16 • Some planetary scientists have suggested that the planet Mars has an electric field somewhat similar to that of the earth, producing a net electric flux of $-3.63 \times 10^{16}\ \mathrm{N \cdot m^2/C}$ at the planet's surface. Calculate: (a) the total electric charge on the planet; (b) the electric field at the planet's surface (refer to the astronomical data in Appendix F or the inside back cover); (c) the charge density on Mars, assuming all the charge is uniformly distributed over the planet's surface.

22.17 •• A very long uniform line of charge has charge per unit length $4.80\ \mu$C/m and lies along the x-axis. A second long uniform line of charge has charge per unit length $-2.40\ \mu$C/m and is parallel to the x-axis at $y = 0.400$ m. What is the net electric field (magnitude and direction) at the following points on the y-axis: (a) $y = 0.200$ m and (b) $y = 0.600$ m?

22.18 •• The electric field 0.400 m from a very long uniform line of charge is 840 N/C. How much charge is contained in a 2.00 cm section of the line?

22.19 •• A hollow, conducting sphere with an outer radius of 0.250 m and an inner radius of 0.200 m has a uniform surface charge density of $+6.37 \times 10^{-6}\ \mathrm{C/m^2}$. A charge of $-0.500\ \mu$C is now introduced at the center of the cavity inside the sphere. (a) What is the new charge density on the outside of the sphere? (b) Calculate the strength of the electric field just outside the sphere. (c) What is the electric flux through a spherical surface just inside the inner surface of the sphere?

22.20 • A solid insulating sphere has total charge Q and radius R. The sphere's charge is distributed uniformly throughout its volume. Let r be the radial distance measured from the center of the sphere. If $E = 800$ N/C at $r = R/2$, what is E at $r = 2R$?

22.21 •• The electric field at a distance of 0.145 m from the surface of a solid insulating sphere with radius 0.355 m is 1750 N/C. (a) Assuming the sphere's charge is uniformly distributed, what is the charge density inside it? (b) Calculate the electric field inside the sphere at a distance of 0.200 m from the center.

22.22 •• A point charge of $-3.00\ \mu$C is located in the center of a spherical cavity of radius 6.50 cm that, in turn, is at the center of an insulating charged solid sphere. The charge density in the solid is $\rho = 7.35 \times 10^{-4}\ \mathrm{C/m^3}$. Calculate the electric field inside the solid at a distance of 9.50 cm from the center of the cavity.

22.23 •• **CP** An electron is released from rest at a distance of 0.300 m from a large insulating sheet of charge that has uniform surface charge density $+2.90 \times 10^{-12}\ \mathrm{C/m^2}$. (a) How much work is done on the electron by the electric field of the sheet as the electron moves from its initial position to a point 0.050 m from the sheet? (b) What is the speed of the electron when it is 0.050 m from the sheet?

22.24 •• Charge Q is distributed uniformly throughout the volume of an insulating sphere of radius $R = 4.00$ cm. At a distance of $r = 8.00$ cm from the center of the sphere, the electric field due to the charge distribution has magnitude $E = 940$ N/C. What are (a) the volume charge density for the sphere and (b) the electric field at a distance of 2.00 cm from the sphere's center?

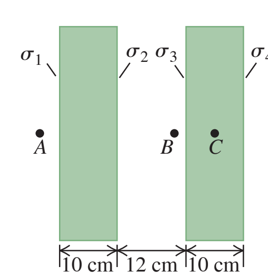
22.25 • A conductor with an inner cavity, like that shown in Fig. 22.23c, carries a total charge of $+5.00$ nC. The charge within the cavity, insulated from the conductor, is -6.00 nC. How much charge is on (a) the inner surface of the conductor and (b) the outer surface of the conductor?

22.26 • A square insulating sheet 80.0 cm on a side is held horizontally. The sheet has 4.50 nC of charge spread uniformly over its area. (a) Estimate the electric field at a point 0.100 mm above the center of the sheet. (b) Estimate the electric field at a point 100 m above the center of the sheet. (c) Would the answers to parts (a) and (b) be different if the sheet were made of a conducting material? Why or why not?

22.27 • An infinitely long cylindrical conductor has radius R and uniform surface charge density σ. (a) In terms of σ and R, what is the charge per unit length λ for the cylinder? (b) In terms of σ, what is the magnitude of the electric field produced by the charged cylinder at a distance $r > R$ from its axis? (c) Express the result of part (b) in terms of λ and show that the electric field outside the cylinder is the same as if all the charge were on the axis. Compare your result to the result for a line of charge in Example 22.6 (Section 22.4).

22.28 • Two very large, nonconducting plastic sheets, each 10.0 cm thick, carry uniform charge densities σ_1, σ_2, σ_3, and σ_4 on their surfaces (**Fig. E22.28**). These surface charge densities have the values $\sigma_1 = -6.00\ \mu\mathrm{C/m^2}$, $\sigma_2 = +5.00\ \mu\mathrm{C/m^2}$, $\sigma_3 = +2.00\ \mu\mathrm{C/m^2}$, and $\sigma_4 = +4.00\ \mu\mathrm{C/m^2}$. Use Gauss's law to find the magnitude and direction of the electric field at the following points, far from the edges of these sheets: (a) point A, 5.00 cm from the left face of the left-hand sheet; (b) point B, 1.25 cm from the inner surface of the right-hand sheet; (c) point C, in the middle of the right-hand sheet.

Figure **E22.28**



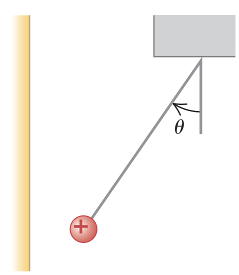
PROBLEMS

22.29 •• **CP** At time $t = 0$ a proton is a distance of 0.360 m from a very large insulating sheet of charge and is moving parallel to the sheet with speed 9.70×10^2 m/s. The sheet has uniform surface charge density $2.34 \times 10^{-9}\ \mathrm{C/m^2}$. What is the speed of the proton at $t = 5.00 \times 10^{-8}$ s?

22.30 •• **CP** A very small object with mass 8.20×10^{-9} kg and positive charge 6.50×10^{-9} C is projected directly toward a very large insulating sheet of positive charge that has uniform surface charge density $5.90 \times 10^{-8}\ \mathrm{C/m^2}$. The object is initially 0.400 m from the sheet. What initial speed must the object have in order for its closest distance of approach to the sheet to be 0.100 m?

22.31 •• **CP** A small sphere with mass 4.00×10^{-6} kg and charge 5.00×10^{-8} C hangs from a thread near a very large, charged insulating sheet (**Fig. P22.31**). The charge density on the surface of the sheet is uniform and equal to $-2.50 \times 10^{-9}\ \mathrm{C/m^2}$. Find the angle of the thread.

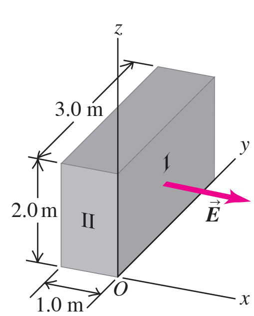
Figure **P22.31**



22.32 •• A cube has sides of length $L = 0.300$ m. One corner is at the origin (Fig. E22.6). The nonuniform electric field is given by $\vec{E} = (-5.00\ \mathrm{N/C \cdot m})x\hat{\imath} + (3.00\ \mathrm{N/C \cdot m})z\hat{k}$. (a) Find the electric flux through each of the six cube faces S_1, S_2, S_3, S_4, S_5, and S_6. (b) Find the total electric charge inside the cube.

22.33 • The electric field $\vec{E}$ in **Fig. P22.33** is everywhere parallel to the x-axis, so the components E_y and E_z are zero. The x-component of the field E_x depends on x but not on y or z. At points in the yz-plane (where $x = 0$), $E_x = 125$ N/C. (a) What is the electric flux through surface I in Fig. P22.33? (b) What is the electric flux through surface II? (c) The volume shown is a small section of a very large insulating slab 1.0 m thick. If there is a total charge of -24.0 nC within the volume shown, what are the magnitude and direction of $\vec{E}$ at the face opposite surface I? (d) Is the electric field produced by charges only within the slab, or is the field also due to charges outside the slab? How can you tell?

Figure **P22.33**



22.34 •• **CALC** In a region of space there is an electric field $\vec{E}$ that is in the z-direction and that has magnitude $E = [964\ \mathrm{N/(C \cdot m)}]x$. Find the flux for this field through a square in the xy-plane at $z = 0$ and with side length 0.350 m. One side of the square is along the $+x$-axis and another side is along the $+y$-axis.

22.35 • Negative charge $-Q$ is distributed uniformly over the surface of a thin spherical insulating shell with radius R. Calculate the force (magnitude and direction) that the shell exerts on a positive point charge q located a distance (a) $r > R$ from the center of the shell (outside the shell); (b) $r < R$ from the center of the shell (inside the shell).

22.36 • A long line carrying a uniform linear charge density $+50.0\ \mu\mathrm{C/m}$ runs parallel to and 10.0 cm from the surface of a large, flat plastic sheet that has a uniform surface charge density of $-100\ \mu\mathrm{C/m^2}$ on one side. Find the location of all points where an α particle would feel no force due to this arrangement of charged objects.

22.37 • **The Coaxial Cable.** A long coaxial cable consists of an inner cylindrical conductor with radius a and an outer coaxial cylinder with inner radius b and outer radius c. The outer cylinder is mounted on insulating supports and has no net charge. The inner cylinder has a uniform positive charge per unit length λ. Calculate the electric field (a) at any point between the cylinders a distance r from the axis and (b) at any point outside the outer cylinder. (c) Graph the magnitude of the electric field as a function of the distance r from the axis of the cable, from $r = 0$ to $r = 2c$. (d) Find the charge per unit length on the inner surface and on the outer surface of the outer cylinder.

22.38 • A very long conducting tube (hollow cylinder) has inner radius a and outer radius b. It carries charge per unit length $+\alpha$, where α is a positive constant with units of C/m. A line of charge lies along the axis of the tube. The line of charge has charge per unit length $+\alpha$. (a) Calculate the electric field in terms of α and the distance r from the axis of the tube for (i) $r < a$; (ii) $a < r < b$; (iii) $r > b$. Show your results in a graph of E as a function of r. (b) What is the charge per unit length on (i) the inner surface of the tube and (ii) the outer surface of the tube?

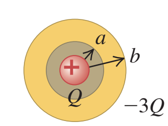
22.39 • A very long, solid cylinder with radius R has positive charge uniformly distributed throughout it, with charge per unit volume ρ. (a) Derive the expression for the electric field inside the volume at a distance r from the axis of the cylinder in terms of the charge density ρ. (b) What is the electric field at a point outside the volume in terms of the charge per unit length λ in the cylinder? (c) Compare the answers to parts (a) and (b) for $r = R$. (d) Graph the electric-field magnitude as a function of r from $r = 0$ to $r = 3R$.

22.40 • **A Sphere in a Sphere.** A solid conducting sphere carrying charge q has radius a. It is inside a concentric hollow conducting sphere with inner radius b and outer radius c. The hollow sphere has no net charge. (a) Derive expressions for the electric-field magnitude in terms of the distance r from the center for the regions $r < a$, $a < r < b$, $b < r < c$, and $r > c$. (b) Graph the magnitude of the electric field as a function of r from $r = 0$ to $r = 2c$. (c) What is the charge on the inner surface of the hollow sphere? (d) On the outer surface? (e) Represent the charge of the small sphere by four plus signs. Sketch the field lines of the system within a spherical volume of radius $2c$.

22.41 • A solid conducting sphere with radius R that carries positive charge Q is concentric with a very thin insulating shell of radius $2R$ that also carries charge Q. The charge Q is distributed uniformly over the insulating shell. (a) Find the electric field (magnitude and direction) in each of the regions $0 < r < R$, $R < r < 2R$, and $r > 2R$. (b) Graph the electric-field magnitude as a function of r.

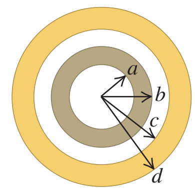
22.42 • A conducting spherical shell with inner radius a and outer radius b has a positive point charge Q located at its center. The total charge on the shell is $-3Q$, and it is insulated from its surroundings (**Fig. P22.42**). (a) Derive expressions for the electric-field magnitude E in terms of the distance r from the center for the regions $r < a$, $a < r < b$, and $r > b$. What is the surface charge density (b) on the inner surface of the conducting shell; (c) on the outer surface of the conducting shell? (d) Sketch the electric field lines and the location of all charges. (e) Graph E as a function of r.

Figure **P22.42**



22.43 • **Concentric Spherical Shells.** A small conducting spherical shell with inner radius a and outer radius b is concentric with a larger conducting spherical shell with inner radius c and outer radius d (**Fig. P22.43**). The inner shell has total charge $+2q$, and the outer shell has charge $+4q$. (a) Calculate the electric field $\vec{E}$ (magnitude and direction) in terms of q and the distance r from the common center of the two shells for (i) $r < a$; (ii) $a < r < b$; (iii) $b < r < c$; (iv) $c < r < d$; (v) $r > d$. Graph the radial component of $\vec{E}$ as a function of r. (b) What is the total charge on the (i) inner surface of the small shell; (ii) outer surface of the small shell; (iii) inner surface of the large shell; (iv) outer surface of the large shell?

Figure **P22.43**



22.44 • Repeat Problem 22.43, but now let the outer shell have charge $-2q$. The inner shell still has charge $+2q$.

22.45 ••• CALC An insulating hollow sphere has inner radius a and outer radius b. Within the insulating material the volume charge density is given by $\rho(r) = \alpha/r$, where α is a positive constant. (a) In terms of α and a, what is the magnitude of the electric field at a distance r from the center of the shell, where $a < r < b$? (b) A point charge q is placed at the center of the hollow space, at $r = 0$. In terms of α and a, what value must q have (sign and magnitude) in order for the electric field to be constant in the region $a < r < b$, and what then is the value of the constant field in this region?

22.46 •• CP **Thomson's Model of the Atom.** Early in the 20th century, a leading model of the structure of the atom was that of English physicist J. J. Thomson (the discoverer of the electron). In Thomson's model, an atom consisted of a sphere of positively charged material in which were embedded negatively charged electrons, like chocolate chips in a ball of cookie dough. Consider such an atom consisting of one electron with mass m and charge $-e$, which may be regarded as a point charge, and a uniformly charged sphere of charge $+e$ and radius R. (a) Explain why the electron's equilibrium position is at the center of the nucleus. (b) In Thomson's model, it was assumed that the positive material provided little or no resistance to the electron's motion. If the electron is displaced from equilibrium by a distance less than R, show that the resulting motion of the electron will be simple harmonic, and calculate the frequency of oscillation. (*Hint:* Review the definition of SHM in Section 14.2. If it can be shown that the net force on the electron is of this form, then it follows that the motion is simple harmonic. Conversely, if the net force on the electron does not follow this form, the motion is not simple harmonic.) (c) By Thomson's time, it was known that excited atoms emit light waves of only certain frequencies. In his model, the frequency of emitted light is the same as the oscillation frequency of the electron(s) in the atom. What radius would a Thomson-model atom need for it to produce red light of frequency 4.57×10^{14} Hz? Compare your answer to the radii of real atoms, which are of the order of 10^{-10} m (see Appendix F). (d) If the electron were displaced from equilibrium by a distance greater than R, would the electron oscillate? Would its motion be simple harmonic? Explain your reasoning. (*Historical note:* In 1910, the atomic nucleus was discovered, proving the Thomson model to be incorrect. An atom's positive charge is not spread over its volume, as Thomson supposed, but is concentrated in the tiny nucleus of radius 10^{-14} to 10^{-15} m.)

22.47 •• CALC A nonuniform, but spherically symmetric, distribution of charge has a charge density $\rho(r)$ given as follows:

$$\rho(r) = \rho_0\left(1 - \frac{r}{R}\right) \qquad \text{for } r \le R$$

$$\rho(r) = 0 \qquad \text{for } r \ge R$$

where $\rho_0 = 3Q/\pi R^3$ is a positive constant. (a) Show that the total charge contained in the charge distribution is Q. (b) Show that the electric field in the region $r \ge R$ is identical to that produced by a point charge Q at $r = 0$. (c) Obtain an expression for the electric field in the region $r \le R$. (d) Graph the electric-field magnitude E as a function of r. (e) Find the value of r at which the electric field is maximum, and find the value of that maximum field.

22.48 •• (a) How many excess electrons must be distributed uniformly within the volume of an isolated plastic sphere 30.0 cm in diameter to produce an electric field of magnitude 1390 N/C just outside the surface of the sphere? (b) What is the electric field at a point 10.0 cm outside the surface of the sphere?

22.49 •• CALC A very long insulating cylinder has radius R and carries positive charge distributed throughout its volume. The charge distribution has cylindrical symmetry but varies with perpendicular distance from the axis of the cylinder. The volume charge density is $\rho(r) = \alpha(1 - r/R)$, where α is a constant with units C/m^3 and r is the perpendicular distance from the center line of the cylinder. Derive an expression, in terms of α and R, for $E(r)$, the electric field as a function of r. Do this for $r < R$ and also for $r > R$. Do your results agree for $r = R$?

22.50 •• CALC A solid insulating sphere has radius R and carries positive charge distributed throughout its volume. The charge distribution has spherical symmetry but varies with radial distance r from the center of the sphere. The volume charge density is $\rho(r) = \rho_0(1 - r/R)$, where ρ_0 is a constant with units of C/m^3. (a) Derive an expression for the electric field as a function of r for $r < R$. (b) Repeat part (a) for $r > R$. (c) At what value of r, in terms of R, does the electric field have its maximum value?

22.51 • CALC **A Nonuniformly Charged Slab.** A slab of insulating material has thickness $2d$ and is oriented so that its faces are parallel to the yz-plane and given by the planes $x = d$ and $x = -d$. The y- and z-dimensions of the slab are very large compared to d, so treat them as infinite. The slab has charge density given by $\rho(x) = \rho_0(x/d)^2$, where ρ_0 is a positive constant. Using Gauss's law, find the electric field due to the slab (magnitude and direction) at all points in space.

22.52 • CALC A nonuniform, but spherically symmetric, distribution of charge has a charge density $\rho(r)$ given as follows:

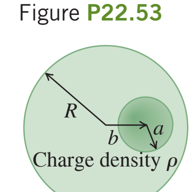
$$\rho(r) = \rho_0\left(1 - \frac{4r}{3R}\right) \qquad \text{for } r \le R$$

$$\rho(r) = 0 \qquad \text{for } r \ge R$$

where ρ_0 is a positive constant. (a) Find the total charge contained in the charge distribution. Obtain an expression for the electric field in the region (b) $r \ge R$; (c) $r \le R$. (d) Graph the electric-field magnitude E as a function of r. (e) Find the value of r at which the electric field is maximum, and find the value of that maximum field.

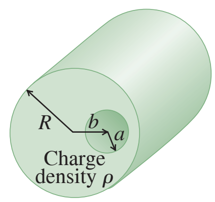
22.53 • (a) An insulating sphere with radius a has a uniform charge density ρ. The sphere is not centered at the origin but at $\vec{r} = \vec{b}$. Show that the electric field inside the sphere is given by $\vec{E} = \rho(\vec{r} - \vec{b})/3\epsilon_0$. (b) An insulating sphere of radius R has a spherical hole of radius a located within its volume and centered a distance b from the center of the sphere, where $a < b < R$ (a cross section of the sphere is shown in **Fig. P22.53**). The solid part of the sphere has a uniform volume charge density ρ. Find the magnitude and direction of the electric field $\vec{E}$ inside the hole, and show that $\vec{E}$ is uniform over the entire hole. [*Hint:* Use the principle of superposition and the result of part (a).]

Figure **P22.53**



22.54 • A very long, solid insulating cylinder has radius R; bored along its entire length is a cylindrical hole with radius a. The axis of the hole is a distance b from the axis of the cylinder, where $a < b < R$ (**Fig. P22.54**). The solid material of the cylinder has a uniform volume charge density ρ. Find the magnitude and direction of the electric field $\vec{E}$ inside the hole, and show that $\vec{E}$ is uniform over the entire hole. (*Hint:* See Problem 22.53.)

Figure **P22.54**



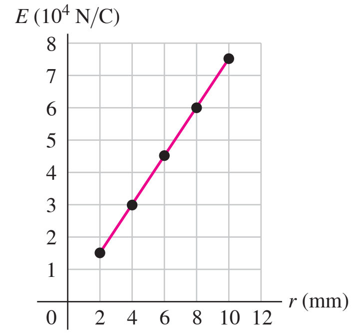
22.55 •• DATA In one experiment the electric field is measured for points at distances r from a uniform line of charge that has charge per unit length λ and length l, where $l \gg r$. In a second experiment the electric field is measured for points at distances r from the center of a uniformly charged insulating sphere that has volume charge density ρ and radius $R = 8.00$ mm, where $r > R$. The results of the two measurements are listed in the table, but you aren't told which set of data applies to which experiment:

	E (10^5 N/C)	
r (cm)	Measurement A	Measurement B
1.00	2.72	5.45
1.50	1.79	2.42
2.00	1.34	1.34
2.50	1.07	0.861
3.00	0.902	0.605
3.50	0.770	0.443
4.00	0.677	0.335

For each set of data, draw two graphs: one for Er^2 versus r and one for Er versus r. (a) Use these graphs to determine which data set, A or B, is for the uniform line of charge and which set is for the uniformly charged sphere. Explain your reasoning. (b) Use the graphs in part (a) to calculate λ for the uniform line of charge and ρ for the uniformly charged sphere.

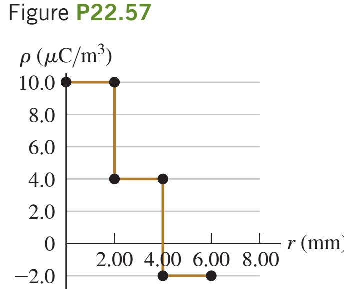
22.56 •• DATA The electric field is measured for points at distances r from the center of a uniformly charged insulating sphere that has volume charge density ρ and radius R, where $r < R$ (**Fig. P22.56**). Calculate ρ.

Figure **P22.56**



22.57 •• DATA The volume charge density ρ for a spherical charge distribution of radius $R = 6.00$ mm is not uniform. **Figure P22.57** shows ρ as a function of the distance r from the center of the distribution. Calculate the electric field at these values of r: (i) 1.00 mm; (ii) 3.00 mm; (iii) 5.00 mm; (iv) 7.00 mm.

Figure **P22.57**



22.58 •• BIO As a honeybee flies, the passing air strips electrons from its hairs, giving the bee a net positive charge. Since flowers are negatively charged, pollen then jumps onto a bee even if the bee does not physically touch the pollen particles. (a) Estimate the diameter of the central disk of a daisy. (b) If a bee has had 75,000 electrons stripped by the air, what is its net charge? (c) If this bee lands at the edge of the daisy's central disk, determine its electric field at the far edge of the disk. Treat the bee as a thin-walled hollow sphere with its net charge distributed uniformly over its surface. (d) A pollen particle requires a force of 10 pN to dislodge from a stamen. Estimate the net charge on a pollen particle at the far end of the disk required for the particle to dislodge and jump to the bee.

22.59 •• CP CALC A very long insulating cylinder with radius R_{cylinder} has nonuniform positive charge density $\rho = (1 - r/R_{\text{cylinder}})\rho_0$, where ρ_0 is constant and r is measured radially from the axis of the cylinder. A particle with charge $-Q$ and mass M orbits the cylinder at a constant distance $R_{\text{orbit}} > R_{\text{cylinder}}$. (a) What is the linear charge density λ of the tube, in terms of R_{cylinder} and ρ_0? (b) Determine the period of the motion in terms of R_{orbit}. (*Hint:* Use Gauss's law to determine the electric field, and therefore the electric force felt by the particle, that acts centripetally.)

22.60 •• CP (a) Show that the component of the electric force normal to any flat surface with a uniform charge density σ is given by $F_\perp = \sigma\,\Phi_E$, where Φ_E is the electric flux through that surface due to an external electric field. (b) An insulating hemisphere with radius R and charge Q distributed uniformly over its flat, circular surface lies above a large plane with uniform charge density σ. The axis of the hemisphere is oriented vertically. For what mass M could the hemisphere remain stationary? (c) If the hemisphere and the plane share the same charge density of 100 μC/m^2 and the hemisphere has a radius of 3.00 cm, what would be its upward acceleration if its mass were 100 g?

CHALLENGE PROBLEMS

22.61 ••• CALC A uniformly charged insulating sphere with radius r and charge $+Q$ lies at the center of a thin-walled hollow cylinder with radius $R > r$ and length $L > 2r$. The cylinder is non-conducting and carries no net charge. (a) Determine the outward electric flux through the rounded "side" of the cylinder, excluding the circular end caps. (*Hint:* Choose a cylindrical coordinate system with the axis of the cylinder as its z-axis and the center of the charged sphere as its origin. Note that an area element on the cylinder has magnitude $dA = 2\pi R dz$.) (b) Determine the electric flux upward through the circular cap at the top of the cylinder. (c) Determine the electric flux downward through the circular cap at the bottom of the cylinder. (d) Add the results from parts (a)–(c) to determine the outward electric flux through the closed cylinder. (e) Show that your result is consistent with Gauss's law.

22.62 ••• CP CALC A region in space contains a total positive charge Q that is distributed spherically such that the volume charge density $\rho(r)$ is given by

$$\begin{aligned} \rho(r) &= 3\alpha r/2R && \text{for } r \le R/2 \\ \rho(r) &= \alpha[1 - (r/R)^2] && \text{for } R/2 \le r \le R \\ \rho(r) &= 0 && \text{for } r \ge R \end{aligned}$$

Here α is a positive constant having units of C/m^3. (a) Determine α in terms of Q and R. (b) Using Gauss's law, derive an expression for the magnitude of the electric field as a function of r. Do this separately for all three regions. Express your answers in terms of Q. (c) What fraction of the total charge is contained within the region $R/2 \le r \le R$? (d) What is the magnitude of $\vec{E}$ at $r = R/2$? (e) If an electron with charge $q' = -e$ is released from rest at any point in any of the three regions, the resulting motion will be oscillatory but not simple harmonic. Why?

MCAT-STYLE PASSAGE PROBLEMS

Space Radiation Shielding. One of the hazards facing humans in space is space radiation: high-energy charged particles emitted by the sun. During a solar flare, the intensity of this radiation can reach lethal levels. One proposed method of protection for astronauts on the surface of the moon or Mars is an array of large, electrically charged spheres placed high above areas where people live and work. The spheres would produce a strong electric field $\vec{E}$ to deflect the charged particles that make up space radiation. The spheres would be similar in construction to a Mylar balloon, with a thin, electrically conducting layer on the outside surface on which a net positive or negative charge would be placed. A typical sphere might be 5 m in diameter.

22.63 Suppose that to repel electrons in the radiation from a solar flare, each sphere must produce an electric field $\vec{E}$ of magnitude 1×10^6 N/C at 25 m from the center of the sphere. What net charge on each sphere is needed? (a) -0.07 C; (b) -8 mC; (c) -80 μC; (d) -1×10^{-20} C.

22.64 What is the magnitude of $\vec{E}$ just outside the surface of such a sphere? (a) 0; (b) 10^6 N/C; (c) 10^7 N/C; (d) 10^8 N/C.

22.65 What is the direction of $\vec{E}$ just outside the surface of such a sphere? (a) Tangent to the surface of the sphere; (b) perpendicular to the surface, pointing toward the sphere; (c) perpendicular to the surface, pointing away from the sphere; (d) there is no electric field just outside the surface.

22.66 Which statement is true about $\vec{E}$ inside a negatively charged sphere as described here? (a) It points from the center of the sphere to the surface and is largest at the center. (b) It points from the surface to the center of the sphere and is largest at the surface. (c) It is zero. (d) It is constant but not zero.

ANSWERS

Chapter Opening Question ?

(iii) The electric field inside a cavity within a conductor is zero, so there would be no electric effect on the child. (See Section 22.5.)

Key Example VARIATION Problems

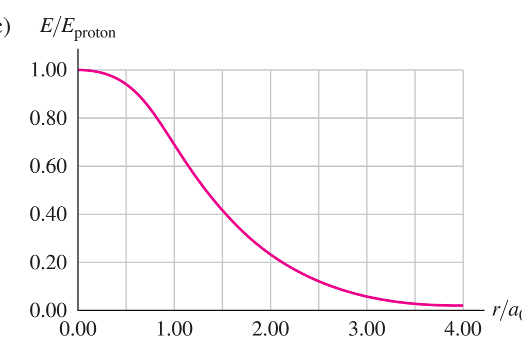
VP22.4.1 (a) $-E_1L^2$ (b) $+E_1L^2$ (c) $-E_2L^2$ (d) $+E_2L^2$ (e) 0 (f) 0 (g) 0
VP22.4.2 (a) $+452\ \text{N}\cdot\text{m}^2/\text{C}$ (b) $-226\ \text{N}\cdot\text{m}^2/\text{C}$ (c) $-226\ \text{N}\cdot\text{m}^2/\text{C}$
VP22.4.3 (a) $+678\ \text{N}\cdot\text{m}^2/\text{C}$ (b) $+113\ \text{N}\cdot\text{m}^2/\text{C}$
VP22.4.4 (a) $+339\ \text{N}\cdot\text{m}^2/\text{C}$ (b) $-565\ \text{N}\cdot\text{m}^2/\text{C}$ (c) $-565\ \text{N}\cdot\text{m}^2/\text{C}$
VP22.10.1 (a) 8.63×10^3 N/C at 4.00 cm, 7.49×10^3 N/C at 6.00 cm (b) 0 at 4.00 cm, 7.49×10^3 N/C at 6.00 cm
VP22.10.2 2.42 μN, attractive
VP22.10.3 1.79×10^{-7} C/m^2, positive
VP22.10.4 5.41×10^{-8} C/m^2
VP22.12.1 (a) -5.00 nC (b) -3.00 nC
VP22.12.2 (a) $+2.00$ nC (b) neither would be affected
VP22.12.3 (a) 0 (b) 2.7×10^{-5} C (c) 2.7×10^{-6} C/m^2
VP22.12.4 (a) $+4.00$ nC (b) -2.00 nC (c) 2.50×10^3 N/C (d) 1.65×10^2 N/C

Bridging Problem

(a) $Q(r) = Qe^{-2r/a_0}[2(r/a_0)^2 + 2(r/a_0) + 1]$

(b) $E = \dfrac{kQe^{-2r/a_0}}{r^2}[2(r/a_0)^2 + 2(r/a_0) + 1]$

(c)



? In one type of welding, electric charge flows between the welding tool and the metal pieces that are to be joined. This produces a glowing arc whose high temperature fuses the pieces together. Why must the tool be held close to the metal pieces? (i) To maximize the potential difference between tool and pieces; (ii) to minimize this potential difference; (iii) to maximize the electric field between tool and pieces; (iv) to minimize this electric field; (v) more than one of these.

23 Electric Potential

LEARNING OUTCOMES

In this chapter, you'll learn...

23.1 How to calculate the electric potential energy of a collection of charges.

23.2 The meaning and significance of electric potential.

23.3 How to calculate the electric potential that a collection of charges produces at a point in space.

23.4 How to use equipotential surfaces to visualize how the electric potential varies in space.

23.5 How to use electric potential to calculate the electric field.

You'll need to review...

7.1–7.4 Conservative forces and potential energy.

21.1–21.6 Electric force and electric fields.

22.4, 22.5 Applications of Gauss's law.

This chapter is about energy associated with electrical interactions. Every time you turn on a light, use a mobile phone, or make toast in a toaster, you are using electrical energy, an indispensable ingredient of our technological society. In Chapters 6 and 7 we introduced the concepts of *work* and *energy* in the context of mechanics; now we'll combine these concepts with what we've learned about electric charge, electric forces, and electric fields. Just as we found for many problems in mechanics, using energy ideas makes it easier to solve a variety of problems in electricity.

When a charged particle moves in an electric field, the field exerts a force that can do *work* on the particle. This work can be expressed in terms of electric potential energy. Just as gravitational potential energy depends on the height of a mass above the earth's surface, electric potential energy depends on the position of the charged particle in the electric field. We'll use a new concept called *electric potential,* or simply *potential* to describe electric potential energy. In circuits, a difference in potential from one point to another is often called *voltage*. The concepts of potential and voltage are crucial to understanding how electric circuits work and have equally important applications to electron beams used in cancer radiotherapy, high-energy particle accelerators, and many other devices.

23.1 ELECTRIC POTENTIAL ENERGY

The concepts of work, potential energy, and conservation of energy proved to be extremely useful in our study of mechanics. In this section we'll show that these concepts are just as useful for understanding and analyzing electrical interactions.

Let's begin by reviewing three essential points from Chapters 6 and 7. First, when a force $\vec{F}$ acts on a particle that moves from point a to point b, the work $W_{a\rightarrow b}$ done by the force is given by a *line integral:*

$$W_{a\rightarrow b} = \int_a^b \vec{F}\cdot d\vec{l} = \int_a^b F\cos\phi\, dl \qquad \text{(work done by a force)} \qquad (23.1)$$

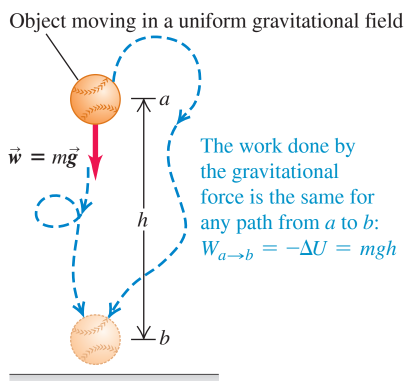
where $d\vec{l}$ is an infinitesimal displacement along the particle's path and ϕ is the angle between $\vec{F}$ and $d\vec{l}$ at each point along the path.

Second, if the force $\vec{F}$ is *conservative*, as we defined the term in Section 7.3, the work done by $\vec{F}$ can always be expressed in terms of a **potential energy** U. When the particle moves from a point where the potential energy is U_a to a point where it is U_b, the change in potential energy is $\Delta U = U_b - U_a$ and

$$\underbrace{W_{a\to b}}_{\text{Work done by a conservative force}} = \overbrace{U_a}^{\text{Potential energy at initial position}} - \underbrace{U_b}_{\text{Potential energy at final position}} = -(U_b - U_a) = \underbrace{-\Delta U}_{\text{Negative of change in potential energy}} \tag{23.2}$$

When $W_{a\to b}$ is positive, U_a is greater than U_b, ΔU is negative, and the potential energy *decreases*. That's what happens when a baseball falls from a high point (a) to a lower point (b) under the influence of the earth's gravity; the force of gravity does positive work, and the gravitational potential energy decreases (**Fig. 23.1**). When a tossed ball is moving upward, the gravitational force does negative work during the ascent, and the potential energy increases.

Figure 23.1 The work done on a baseball moving in a uniform gravitational field.



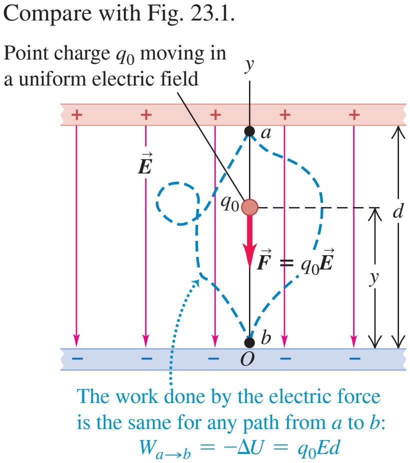
Third, the work–energy theorem says that the change in kinetic energy $\Delta K = K_b - K_a$ during a displacement equals the *total* work done on the particle. If only conservative forces do work, then Eq. (23.2) gives the total work, and $K_b - K_a = -(U_b - U_a)$. We usually write this as

$$K_a + U_a = K_b + U_b \tag{23.3}$$

That is, the total mechanical energy (kinetic plus potential) is *conserved* under these circumstances.

Electric Potential Energy in a Uniform Field

Figure 23.2 The work done on a point charge moving in a uniform electric field. Compare with Fig. 23.1.



Let's look at an electrical example of these concepts. In **Fig. 23.2** a pair of large, charged, parallel metal plates sets up a uniform, downward electric field with magnitude E. The field exerts a downward force with magnitude $F = q_0E$ on a positive test charge q_0. As the charge moves downward a distance d from point a to point b, the force on the test charge is constant and independent of its location. So the work done by the electric field is the product of the force magnitude and the component of displacement in the (downward) direction of the force:

$$W_{a\to b} = Fd = q_0Ed \tag{23.4}$$

This work is positive, since the force is in the same direction as the net displacement of the test charge.

The y-component of the electric force, $F_y = -q_0E$, is constant, and there is no x- or z-component. This is exactly analogous to the gravitational force on a mass m near the earth's surface; for this force, there is a constant y-component $F_y = -mg$ and the x- and z-components are zero. Because of this analogy, we can conclude that the force exerted on q_0 by the uniform electric field in Fig. 23.2 is *conservative,* just as is the gravitational force. This means that the work $W_{a\to b}$ done by the field is independent of the path the particle takes from a to b. We can represent this work with a *potential-energy* function U, just as we did for gravitational potential energy in Section 7.1. The potential energy for the gravitational force $F_y = -mg$ was $U = mgy$; hence the potential energy for the electric force $F_y = -q_0E$ is

$$U = q_0Ey \tag{23.5}$$

When the test charge moves from height y_a to height y_b, the work done on the charge by the field is given by

$$W_{a\to b} = -\Delta U = -(U_b - U_a) = -(q_0Ey_b - q_0Ey_a) = q_0E(y_a - y_b) \tag{23.6}$$

Figure 23.3 A positive charge moving (a) in the direction of the electric field $\vec{E}$ and (b) in the direction opposite $\vec{E}$.

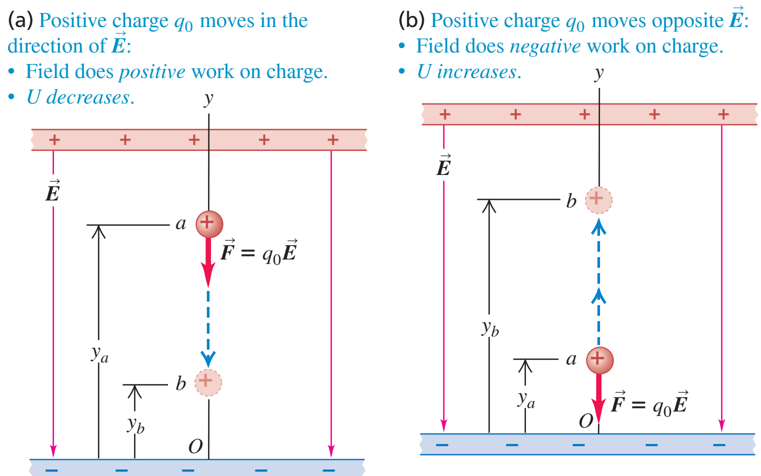
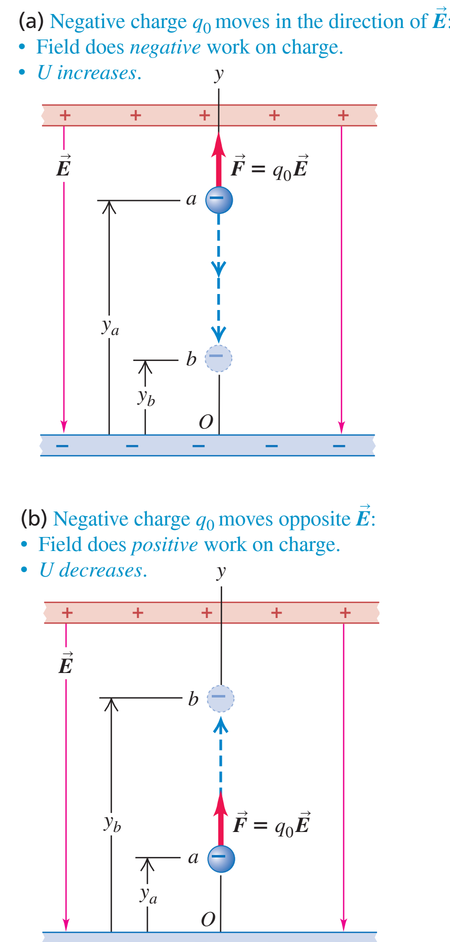


Figure 23.4 A negative charge moving (a) in the direction of the electric field $\vec{E}$ and (b) in the direction opposite $\vec{E}$. Compare with Fig. 23.3.



When y_a is greater than y_b (**Fig. 23.3a**), the positive test charge q_0 moves downward, in the same direction as $\vec{E}$; the displacement is in the same direction as the force $\vec{F} = q_0\vec{E}$, so the field does positive work and U decreases. [In particular, if $y_a - y_b = d$ as in Fig. 23.2, Eq. (23.6) gives $W_{a \to b} = q_0Ed$, in agreement with Eq. (23.4).] When y_a is less than y_b (Fig. 23.3b), the positive test charge q_0 moves upward, in the opposite direction to $\vec{E}$; the displacement is opposite the force, the field does negative work, and U increases.

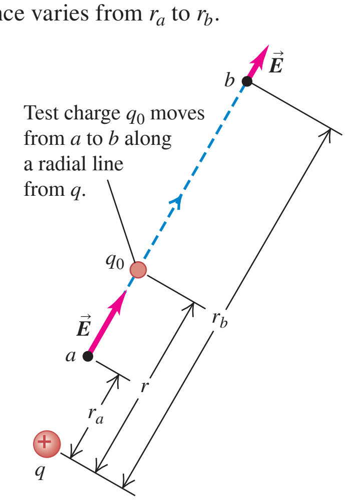
If the test charge q_0 is *negative,* the potential energy increases when it moves with the field and decreases when it moves against the field (**Fig. 23.4**).

Whether the test charge is positive or negative, the following general rules apply: *U increases* if the test charge q_0 moves in the direction *opposite* the electric force $\vec{F} = q_0\vec{E}$ (Figs. 23.3b and 23.4a); *U decreases* if q_0 moves in the *same* direction as $\vec{F} = q_0\vec{E}$ (Figs. 23.3a and 23.4b). This is the same behavior as for gravitational potential energy, which increases if a mass m moves upward (opposite the direction of the gravitational force) and decreases if m moves downward (in the same direction as the gravitational force).

CAUTION **Electric potential energy** The relationship between electric potential energy change and motion in an electric field is an important one that we'll use often, but that takes some effort to understand. Studying the preceding paragraph as well as Figs. 23.3 and Fig. 23.4 carefully now will help you tremendously later!

Electric Potential Energy of Two Point Charges

Figure 23.5 Test charge q_0 moves along a straight line extending radially from charge q. As it moves from a to b, the distance varies from r_a to r_b.



The idea of electric potential energy isn't restricted to the special case of a uniform electric field. Indeed, we can apply this concept to a point charge in *any* electric field caused by a static charge distribution. Recall from Chapter 21 that we can represent any charge distribution as a collection of point charges. Therefore it's useful to calculate the work done on a test charge q_0 moving in the electric field caused by a single, stationary point charge q.

We'll consider first a displacement along the *radial* line in **Fig. 23.5**. The force on q_0 is given by Coulomb's law, and its radial component is

$$F_r = \frac{1}{4\pi\epsilon_0}\frac{qq_0}{r^2} \tag{23.7}$$

If q and q_0 have the same sign $(+ \text{ or } -)$, the force is repulsive and F_r is positive; if the two charges have opposite signs, the force is attractive and F_r is negative. The force is *not*

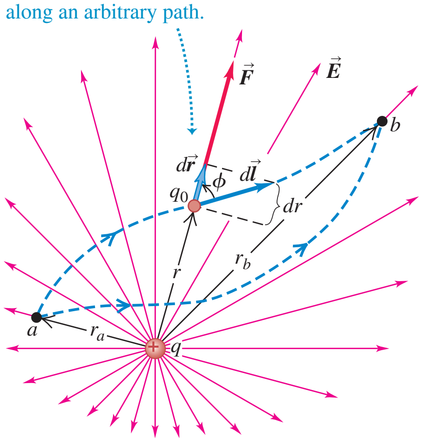


Figure 23.6 The work done on charge q_0 by the electric field of charge q does not depend on the path taken, but only on the distances r_a and r_b.

constant during the displacement, and we must integrate to calculate the work $W_{a \to b}$ done on q_0 by this force as q_0 moves from a to b:

$$W_{a \to b} = \int_{r_a}^{r_b} F_r\, dr = \int_{r_a}^{r_b} \frac{1}{4\pi\epsilon_0} \frac{qq_0}{r^2} dr = \frac{qq_0}{4\pi\epsilon_0}\left(\frac{1}{r_a} - \frac{1}{r_b}\right) \tag{23.8}$$

The work done by the electric force for this path depends on only the endpoints.

Now let's consider a more general displacement (**Fig. 23.6**) in which a and b do *not* lie on the same radial line. From Eq. (23.1) the work done on q_0 during this displacement is given by

$$W_{a \to b} = \int_{r_a}^{r_b} F \cos\phi\, dl = \int_{r_a}^{r_b} \frac{1}{4\pi\epsilon_0} \frac{qq_0}{r^2} \cos\phi\, dl$$

But Fig. 23.6 shows that $\cos\phi\, dl = dr$. That is, the work done during a small displacement $d\vec{l}$ depends only on the change dr in the distance r between the charges, which is the *radial component* of the displacement. Thus Eq. (23.8) is valid even for this more general displacement; the work done on q_0 by the electric field $\vec{E}$ produced by q depends only on r_a and r_b, not on the details of the path. Also, if q_0 returns to its starting point a by a different path, the total work done in the round-trip displacement is zero (the integral in Eq. (23.8) is from r_a back to r_a). These are the needed characteristics for a conservative force, as we defined it in Section 7.3. Thus the force on q_0 is a *conservative* force.

We see that Eqs. (23.2) and (23.8) are consistent if we define the potential energy to be $U_a = qq_0/4\pi\epsilon_0 r_a$ when q_0 is a distance r_a from q, and to be $U_b = qq_0/4\pi\epsilon_0 r_b$ when q_0 is a distance r_b from q. Thus

Electric potential energy of two point charges

$$U = \frac{1}{4\pi\epsilon_0} \frac{qq_0}{r} \tag{23.9}$$

Values of two charges (qq_0); Distance between two charges (r); Electric constant (ϵ_0)

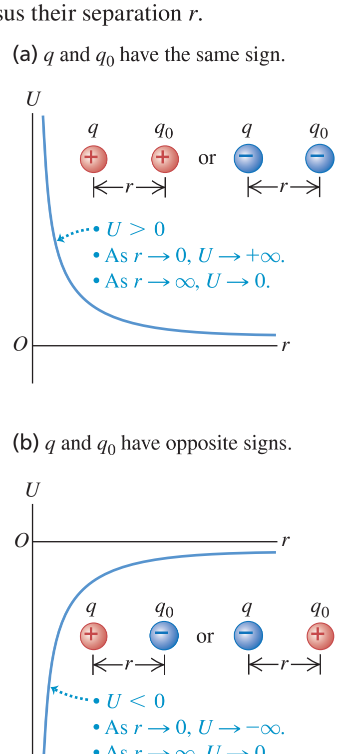


Figure 23.7 Graphs of the potential energy U of two point charges q and q_0 versus their separation r.

Equation (23.9) is valid no matter what the signs of the charges q and q_0. The potential energy is positive if the charges q and q_0 have the same sign (**Fig. 23.7a**) and negative if they have opposite signs (Fig. 23.7b).

CAUTION **Electric potential energy vs. electric force** Don't confuse Eq. (23.9) for the potential energy of two point charges with Eq. (23.7) for the radial component of the electric force that one charge exerts on the other. Potential energy U is proportional to $1/r$, while the force component F_r is proportional to $1/r^2$.

Potential energy is always defined relative to some reference point where $U = 0$. In Eq. (23.9), U is zero when q and q_0 are infinitely far apart and $r = \infty$. Therefore U represents the work that would be done on the test charge q_0 by the field of q if q_0 moved from an initial distance r to infinity. If q and q_0 have the same sign, the interaction is repulsive, this work is positive, and U is positive at any finite separation (Fig. 23.7a). If the charges have opposite signs, the interaction is attractive, the work done is negative, and U is negative (Fig. 23.7b).

We emphasize that the potential energy U given by Eq. (23.9) is a *shared* property of the two charges. If the distance between q and q_0 is changed from r_a to r_b, the change in potential energy is the same whether q is held fixed and q_0 is moved or q_0 is held fixed and q is moved. For this reason, we never use the phrase "the electric potential energy *of* a point charge." (Likewise, if a mass m is at a height h above the earth's surface, the gravitational potential energy is a shared property of the mass m and the earth. We emphasized this in Sections 7.1 and 13.3.)

Equation (23.9) also holds if the charge q_0 is outside a spherically symmetric charge *distribution* with total charge q; the distance r is from q_0 to the center of the distribution. That's because Gauss's law tells us that the electric field outside such a distribution is the same as if all of its charge q were concentrated at its center (see Example 22.9 in Section 22.4).

EXAMPLE 23.1 Conservation of energy with electric forces

WITH VARIATION PROBLEMS

A positron (the electron's antiparticle) has mass 9.11×10^{-31} kg and charge $q_0 = +e = +1.60 \times 10^{-19}$ C. Suppose a positron moves in the vicinity of an α (alpha) particle, which has charge $q = +2e = 3.20 \times 10^{-19}$ C and mass 6.64×10^{-27} kg. The α particle's mass is more than 7000 times that of the positron, so we assume that the α particle remains at rest. When the positron is 1.00×10^{-10} m from the α particle, it is moving directly away from the α particle at 3.00×10^6 m/s. (a) What is the positron's speed when the particles are 2.00×10^{-10} m apart? (b) What is the positron's speed when it is very far from the α particle? (c) Suppose the initial conditions are the same but the moving particle is an electron (with the same mass as the positron but charge $q_0 = -e$). Describe the subsequent motion.

IDENTIFY and SET UP The electric force between a positron (or an electron) and an α particle is conservative, so the total mechanical energy (kinetic plus potential) is conserved. Equation (23.9) gives the potential energy U at any separation r: The potential-energy function for parts (a) and (b) looks like that of Fig. 23.7a, and the function for part (c) looks like that of Fig. 23.7b. We are given the positron speed $v_a = 3.00 \times 10^6$ m/s when the separation between the particles is $r_a = 1.00 \times 10^{-10}$ m. In parts (a) and (b) we use Eqs. (23.3) and (23.9) to find the speed for $r = r_b = 2.00 \times 10^{-10}$ m and $r = r_c \rightarrow \infty$, respectively. In part (c) we replace the positron with an electron and reconsider the problem.

EXECUTE (a) Both particles have positive charge, so the positron speeds up as it moves away from the α particle. From the energy-conservation equation, Eq. (23.3), the final kinetic energy is

$$K_b = \tfrac{1}{2}mv_b^2 = K_a + U_a - U_b$$

In this expression,

$$K_a = \tfrac{1}{2}mv_a^2 = \tfrac{1}{2}(9.11 \times 10^{-31}\ \text{kg})(3.00 \times 10^6\ \text{m/s})^2$$
$$= 4.10 \times 10^{-18}\ \text{J}$$
$$U_a = \frac{1}{4\pi\epsilon_0}\frac{qq_0}{r_a} = (9.0 \times 10^9\ \text{N}\cdot\text{m}^2/\text{C}^2)$$
$$\times \frac{(3.20 \times 10^{-19}\ \text{C})(1.60 \times 10^{-19}\ \text{C})}{1.00 \times 10^{-10}\ \text{m}}$$
$$= 4.61 \times 10^{-18}\ \text{J}$$
$$U_b = \frac{1}{4\pi\epsilon_0}\frac{qq_0}{r_b} = 2.30 \times 10^{-18}\ \text{J}$$

Hence the positron kinetic energy and speed at $r = r_b = 2.00 \times 10^{-10}$ m are

$$K_b = \tfrac{1}{2}mv_b^2$$
$$= (4.10 \times 10^{-18}\ \text{J}) + (4.61 \times 10^{-18}\ \text{J}) - (2.30 \times 10^{-18}\ \text{J})$$
$$= 6.41 \times 10^{-18}\ \text{J}$$
$$v_b = \sqrt{\frac{2K_b}{m}} = \sqrt{\frac{2(6.41 \times 10^{-18}\ \text{J})}{9.11 \times 10^{-31}\ \text{kg}}}$$
$$= 3.8 \times 10^6\ \text{m/s}$$

(b) When the positron and α particle are very far apart so that $r = r_c \rightarrow \infty$, the final potential energy U_c approaches zero. Again from energy conservation, the final kinetic energy and speed of the positron in this case are

$$K_c = K_a + U_a - U_c = (4.10 \times 10^{-18}\ \text{J}) + (4.61 \times 10^{-18}\ \text{J}) - 0$$
$$= 8.71 \times 10^{-18}\ \text{J}$$
$$v_c = \sqrt{\frac{2K_c}{m}} = \sqrt{\frac{2(8.71 \times 10^{-18}\ \text{J})}{9.11 \times 10^{-31}\ \text{kg}}} = 4.4 \times 10^6\ \text{m/s}$$

(c) The electron and α particle have opposite charges, so the force is attractive and the electron slows down as it moves away. Changing the moving particle's sign from $+e$ to $-e$ means that the initial potential energy is now $U_a = -4.61 \times 10^{-18}$ J, which makes the total mechanical energy *negative*:

$$K_a + U_a = (4.10 \times 10^{-18}\ \text{J}) - (4.61 \times 10^{-18}\ \text{J})$$
$$= -0.51 \times 10^{-18}\ \text{J}$$

The total mechanical energy would have to be positive for the electron to move infinitely far away from the α particle. Like a rock thrown upward at low speed from the earth's surface, it will reach a maximum separation $r = r_d$ from the α particle before reversing direction. At this point its speed and its kinetic energy K_d are zero, so at separation r_d we have

$$U_d = K_a + U_a - K_d = (-0.51 \times 10^{-18}\ \text{J}) - 0$$
$$U_d = \frac{1}{4\pi\epsilon_0}\frac{qq_0}{r_d} = -0.51 \times 10^{-18}\ \text{J}$$
$$r_d = \frac{1}{U_d}\frac{qq_0}{4\pi\epsilon_0}$$
$$= \frac{(9.0 \times 10^9\ \text{N}\cdot\text{m}^2/\text{C}^2)}{-0.51 \times 10^{-18}\ \text{J}}(3.20 \times 10^{-19}\ \text{C})(-1.60 \times 10^{-19}\ \text{C})$$
$$= 9.0 \times 10^{-10}\ \text{m}$$

For $r_b = 2.00 \times 10^{-10}$ m we have $U_b = -2.30 \times 10^{-18}$ J, so the electron kinetic energy and speed at this point are

$$K_b = \tfrac{1}{2}mv_b^2$$
$$= (4.10 \times 10^{-18}\ \text{J}) + (-4.61 \times 10^{-18}\ \text{J}) - (-2.30 \times 10^{-18}\ \text{J})$$
$$= 1.79 \times 10^{-18}\ \text{J}$$
$$v_b = \sqrt{\frac{2K_b}{m}} = \sqrt{\frac{2(1.79 \times 10^{-18}\ \text{J})}{9.11 \times 10^{-31}\ \text{kg}}} = 2.0 \times 10^6\ \text{m/s}$$

EVALUATE Both particles behave as expected as they move away from the α particle: The positron speeds up, and the electron slows down and eventually turns around. How fast would an electron have to be moving at $r_a = 1.00 \times 10^{-10}$ m to travel infinitely far from the α particle? (*Hint:* See Example 13.5 in Section 13.3.)

KEYCONCEPT The electrostatic forces that two charged particles exert on each other are conservative. The associated electric potential energy U is proportional to the product of the two charges and inversely proportional to the distance r between them. As r approaches infinity, U approaches zero.

Electric Potential Energy with Several Point Charges

Suppose the electric field $\vec{E}$ in which charge q_0 moves is caused by *several* point charges $q_1, q_2, q_3, \ldots$ at distances $r_1, r_2, r_3, \ldots$ from q_0, as in **Fig. 23.8**. For example, q_0 could be a positive ion moving in the presence of other ions (**Fig. 23.9**). The total electric field at each point is the *vector sum* of the fields due to the individual charges, and the total work done on q_0 during any displacement is the sum of the contributions from the individual charges. From Eq. (23.9) we conclude that the potential energy associated with the test charge q_0 at point a in Fig. 23.8 is the *algebraic* sum (*not* a vector sum):

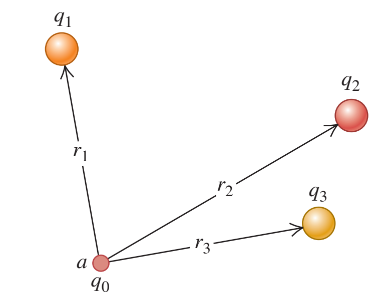


Figure **23.8** The potential energy associated with a charge q_0 at point a depends on the other charges q_1, q_2, and q_3 and on their distances r_1, r_2, and r_3 from point a.

Electric potential energy of point charge q_0 and collection of charges $q_1, q_2, q_3, \ldots$

$$U = \frac{q_0}{4\pi\epsilon_0}\left(\frac{q_1}{r_1} + \frac{q_2}{r_2} + \frac{q_3}{r_3} + \cdots\right) = \frac{q_0}{4\pi\epsilon_0}\sum_i \frac{q_i}{r_i} \qquad (23.10)$$

Electric constant

Distances from q_0 to $q_1, q_2, q_3, \ldots$

When q_0 is at a different point b, the potential energy is given by the same expression, but $r_1, r_2, \ldots$ are the distances from $q_1, q_2, \ldots$ to point b. The work done on charge q_0 when it moves from a to b along any path is equal to the difference $U_a - U_b$ between the potential energies when q_0 is at a and at b.

We can represent *any* charge distribution as a collection of point charges, so Eq. (23.10) shows that we can always find a potential-energy function for *any* static electric field. It follows that **for every electric field due to a static charge distribution, the force exerted by that field is conservative.**

Equations (23.9) and (23.10) define U to be zero when distances $r_1, r_2, \ldots$ are infinite—that is, when the test charge q_0 is very far away from all the charges that produce the field. As with any potential-energy function, the point where $U = 0$ is arbitrary; we can always add a constant to make U equal zero at any point we choose. In electrostatics problems it's usually simplest to choose this point to be at infinity. When we analyze electric circuits in Chapters 25 and 26, other choices will be more convenient.

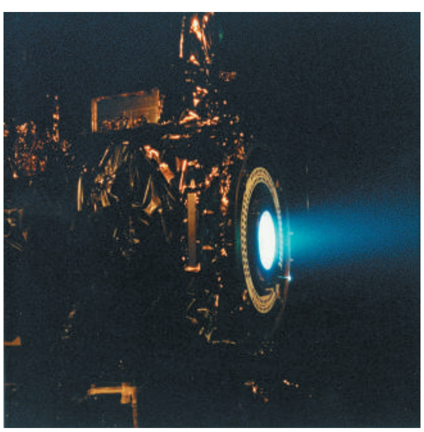


Figure **23.9** This ion engine for spacecraft uses electric forces to eject a stream of positive xenon ions (Xe^+) at speeds in excess of 30 km/s. The thrust produced is very low (about 0.09 newton) but can be maintained continuously for days, in contrast to chemical rockets, which produce a large thrust for a short time (see Fig. 8.34). Such ion engines have been used for maneuvering interplanetary spacecraft.

Equation (23.10) gives the potential energy associated with the presence of the test charge q_0 in the $\vec{E}$ field produced by $q_1, q_2, q_3, \ldots$. But there is also potential energy involved in assembling these charges. If we start with charges $q_1, q_2, q_3, \ldots$ all separated from each other by infinite distances and then bring them together so that the distance between q_i and q_j is r_{ij}, the *total* potential energy U is the sum of the potential energies of interaction for each pair of charges. We can write this as

$$U = \frac{1}{4\pi\epsilon_0}\sum_{i<j}\frac{q_i q_j}{r_{ij}} \qquad (23.11)$$

This sum extends over all *pairs* of charges; we don't let $i = j$ (because that would be an interaction of a charge with itself), and we include only terms with $i < j$ to make sure that we count each pair only once. Thus, to account for the interaction between q_3 and q_4, we include a term with $i = 3$ and $j = 4$ but not a term with $i = 4$ and $j = 3$.

Interpreting Electric Potential Energy

As a final comment, here are two viewpoints on electric potential energy. We have defined it in terms of the work done *by the electric field* on a charged particle moving in the field, just as in Chapter 7 we defined potential energy in terms of the work done by gravity or by a spring. When a particle moves from point a to point b, the work done on it by the electric field is $W_{a\to b} = U_a - U_b$. Thus the potential-energy difference $U_a - U_b$ equals *the work that is done by the electric force when the particle moves from a to b*. When U_a is greater than U_b, the field does positive work on the particle as it "falls" from a point of higher potential energy (a) to a point of lower potential energy (b).

An alternative but equivalent viewpoint is to consider how much work we would have to do to "raise" a particle from a point b where the potential energy is U_b to a point a where it has a greater value U_a (pushing two positive charges closer together, for example). To move the particle slowly (so as not to give it any kinetic energy), we need to exert an additional external force $\vec{F}_{\text{ext}}$ that is equal and opposite to the electric-field force and does positive work. The potential-energy difference $U_a - U_b$ is then defined as *the work that must be done by an external force to move the particle slowly from b to a against the electric force*. Because $\vec{F}_{\text{ext}}$ is the negative of the electric-field force and the displacement is in the opposite direction, this definition of the potential difference $U_a - U_b$ is equivalent to that given above. This alternative viewpoint also works if U_a is less than U_b, corresponding to "lowering" the particle; an example is moving two positive charges away from each other. In this case, $U_a - U_b$ is again equal to the work done by the external force, but now this work is negative.

We'll use both of these viewpoints in the next section to interpret what is meant by electric *potential,* or potential energy per unit charge.

EXAMPLE 23.2 A system of point charges

WITH VARIATION PROBLEMS

Two point charges are at fixed positions on the x-axis, $q_1 = -e$ at $x = 0$ and $q_2 = +e$ at $x = a$. (a) Find the work that must be done by an external force to bring a third point charge $q_3 = +e$ from infinity to $x = 2a$. (b) Find the total potential energy of the system of three charges.

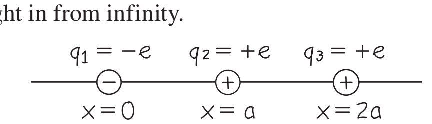
IDENTIFY and SET UP **Figure 23.10** shows the final arrangement of the three charges. In part (a) we need to find the work W that must be done on q_3 by an external force $\vec{F}_{\text{ext}}$ to bring q_3 in from infinity to $x = 2a$. We do this by using Eq. (23.10) to find the potential energy associated with q_3 in the presence of q_1 and q_2. In part (b) we use Eq. (23.11), the expression for the potential energy of a collection of point charges, to find the total potential energy of the system.

EXECUTE (a) The work W equals the difference between (i) the potential energy U associated with q_3 when it is at $x = 2a$ and (ii) the potential energy when it is infinitely far away. The second of these is zero, so the work required is equal to U. The distances between the charges are $r_{13} = 2a$ and $r_{23} = a$, so from Eq. (23.10),

$$W = U = \frac{q_3}{4\pi\epsilon_0}\left(\frac{q_1}{r_{13}} + \frac{q_2}{r_{23}}\right) = \frac{+e}{4\pi\epsilon_0}\left(\frac{-e}{2a} + \frac{+e}{a}\right) = \frac{+e^2}{8\pi\epsilon_0 a}$$

This is positive, just as we should expect. If we bring q_3 in from infinity along the $+x$-axis, it is attracted by q_1 but is repelled more strongly by q_2. Hence we must do positive work to push q_3 to the position at $x = 2a$.

Figure **23.10** Our sketch of the situation after the third charge has been brought in from infinity.



(b) From Eq. (23.11), the total potential energy of the three-charge system is

$$U = \frac{1}{4\pi\epsilon_0}\sum_{i<j}\frac{q_i q_j}{r_{ij}} = \frac{1}{4\pi\epsilon_0}\left(\frac{q_1 q_2}{r_{12}} + \frac{q_1 q_3}{r_{13}} + \frac{q_2 q_3}{r_{23}}\right)$$

$$= \frac{1}{4\pi\epsilon_0}\left[\frac{(-e)(e)}{a} + \frac{(-e)(e)}{2a} + \frac{(e)(e)}{a}\right] = \frac{-e^2}{8\pi\epsilon_0 a}$$

EVALUATE Our negative result in part (b) means that the system has lower potential energy than it would if the three charges were infinitely far apart. An external force would have to do *negative* work to bring the three charges from infinity to assemble this entire arrangement and would have to do *positive* work to move the three charges back to infinity.

KEYCONCEPT To find the total potential energy of a system of charges, sum the electric potential energies of all the distinct pairs of charges in the system.

TEST YOUR UNDERSTANDING OF SECTION 23.1 Consider the system of three point charges in Example 21.4 (Section 21.3) and shown in Fig. 21.14. (a) What is the sign of the total potential energy of this system? (i) Positive; (ii) negative; (iii) zero. (b) What is the sign of the total amount of work you would have to do to move these charges infinitely far from each other? (i) Positive; (ii) negative; (iii) zero.

ANSWER

(a) **(i)**, (b) **(ii)** The three charges q_1, q_2, and q_3 are all positive, so all three of the terms in the sum in Eq. (23.11)—q_1q_2/r_{12}, q_1q_3/r_{13}, and q_2q_3/r_{23}—are positive. Hence the total electric potential energy U is positive. This means that it would take positive work to bring the three charges from infinity to the positions shown in Fig. 21.14, and hence *negative* work to move the three charges from these positions back to infinity.

23.2 ELECTRIC POTENTIAL

In Section 23.1 we looked at the potential energy U associated with a test charge q_0 in an electric field. Now we want to describe this potential energy on a "per unit charge" basis, just as electric field describes the force per unit charge on a charged particle in the field. This leads us to the concept of *electric potential,* often called simply *potential*. This concept is very useful in calculations involving energies of charged particles. It also facilitates many electric-field calculations because electric potential is closely related to the electric field $\vec{E}$. When we need to determine an electric field, it is often easier to determine the potential first and then find the field from it.

Potential is *potential energy per unit charge*. We define the potential V at any point in an electric field as the potential energy U *per unit charge* associated with a test charge q_0 at that point:

$$V = \frac{U}{q_0} \quad \text{or} \quad U = q_0 V \tag{23.12}$$

Potential energy and charge are both scalars, so potential is a scalar. From Eq. (23.12) its units are the units of energy divided by those of charge. The SI unit of potential, called one **volt** (1 V) in honor of the Italian electrical experimenter Alessandro Volta (1745–1827), equals 1 joule per coulomb:

$$1\text{ V} = 1\text{ volt} = 1\text{ J/C} = 1\text{ joule/coulomb}$$

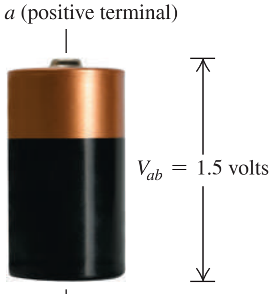
Let's put Eq. (23.2), which equates the work done by the electric force during a displacement from a to b to the quantity $-\Delta U = -(U_b - U_a)$, on a "work per unit charge" basis. We divide this equation by q_0, obtaining

$$\frac{W_{a\to b}}{q_0} = -\frac{\Delta U}{q_0} = -\left(\frac{U_b}{q_0} - \frac{U_a}{q_0}\right) = -(V_b - V_a) = V_a - V_b \tag{23.13}$$

where $V_a = U_a/q_0$ is the potential energy per unit charge at point a and similarly for V_b. We call V_a and V_b the *potential at point a* and *potential at point b*, respectively. Thus the work done per unit charge by the electric force when a charged object moves from a to b is equal to the potential at a minus the potential at b.

The difference $V_a - V_b$ is called the *potential of a with respect to b*; we sometimes abbreviate this difference as $V_{ab} = V_a - V_b$ (note the order of the subscripts). This is often called the potential difference between a and b, but that's ambiguous unless we specify which is the reference point. In electric circuits, which we'll analyze in later chapters, the potential difference between two points is often called **voltage** (**Fig. 23.11**). Equation (23.13) then states: **V_{ab}, the potential (in V) of a with respect to b, equals the work (in J) done by the electric force when a UNIT (1 C) charge moves from a to b.**

Figure 23.11 The voltage of this battery equals the difference in potential $V_{ab} = V_a - V_b$ between its positive terminal (point a) and its negative terminal (point b).



Another way to interpret the potential difference V_{ab} in Eq. (23.13) is to use the alternative viewpoint mentioned at the end of Section 23.1. In that viewpoint, $U_a - U_b$ is the amount of work that must be done by an *external* force to move a particle of charge q_0 slowly from b to a against the electric force. The work that must be done *per unit charge* by the external force is then $(U_a - U_b)/q_0 = V_a - V_b = V_{ab}$. In other words: **$V_{ab}$, the potential (in V) of a with respect to b, equals the work (in J) that must be done to move a UNIT (1 C) charge slowly from b to a against the electric force.**

An instrument that measures the difference of potential between two points is called a *voltmeter*. (In Chapter 26 we'll discuss how these devices work.) Voltmeters that can measure a potential difference of 1 μV are common, and sensitivities down to 10^{-12} V can be attained.

Calculating Electric Potential

To find the potential V due to a single point charge q, we divide Eq. (23.9) by q_0:

Electric potential due to a point charge

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{r} \qquad (23.14)$$

Value of point charge

Distance from point charge to where potential is measured

Electric constant

If q is positive, the potential that it produces is positive at all points; if q is negative, it produces a potential that is negative everywhere. In either case, V is equal to zero at $r = \infty$, an infinite distance from the point charge. Note that potential, like electric field, is independent of the test charge q_0 that we use to define it.

Similarly, we divide Eq. (23.10) by q_0 to find the potential due to a collection of point charges:

Electric potential due to a collection of point charges

$$V = \frac{1}{4\pi\epsilon_0}\sum_i \frac{q_i}{r_i} \qquad (23.15)$$

Value of ith point charge

Distance from ith point charge to where potential is measured

Electric constant

Just as the electric field due to a collection of point charges is the *vector* sum of the fields produced by each charge, the electric potential due to a collection of point charges is the *scalar* sum of the potentials due to each charge. When we have a continuous distribution of charge along a line, over a surface, or through a volume, we divide the charge into elements dq, and the sum in Eq. (23.15) becomes an integral:

Integral over charge distribution

Electric potential due to a continuous distribution of charge

$$V = \frac{1}{4\pi\epsilon_0}\int \frac{dq}{r} \qquad (23.16)$$

Charge element

Distance from charge element to where potential is measured

Electric constant

We'll work out several examples of such cases. The potential defined by Eqs. (23.15) and (23.16) is zero at points that are infinitely far away from *all* the charges. Later we'll encounter cases in which the charge distribution itself extends to infinity. We'll find that in such cases we cannot set $V = 0$ at infinity, and we'll need to exercise care in using and interpreting Eqs. (23.15) and (23.16).

CAUTION **What is electric potential?** Before getting too involved in the details of how to calculate electric potential, remind yourself what potential is. The electric *potential* at a certain point is the potential energy per *unit* charge placed at that point. That's why potential is measured in joules per coulomb, or volts. Keep in mind, too, that there doesn't have to be a charge at a given point for a potential V to exist at that point. (In the same way, an electric field can exist at a given point even if there's no charge there to respond to it.)

Finding Electric Potential from Electric Field

When we are given a collection of point charges, Eq. (23.15) is usually the easiest way to calculate the potential V. But in some problems in which the electric field is known or can be found easily, it is easier to determine V from $\vec{E}$. The force $\vec{F}$ on a test charge q_0 can be written as $\vec{F} = q_0\vec{E}$, so from Eq. (23.1) the work done by the electric force as the test charge moves from a to b is given by

$$W_{a\to b} = \int_a^b \vec{F}\cdot d\vec{l} = \int_a^b q_0\vec{E}\cdot d\vec{l}$$

BIO APPLICATION Electrocardiography The electrodes used in an electrocardiogram—EKG or ECG for short—measure the potential differences (typically no greater than 1 mV = 10^{-3} V) between different parts of the patient's skin. These are indicative of the potential differences between regions of the heart, and so provide a sensitive way to detect any abnormalities in the electrical activity that drives cardiac function.

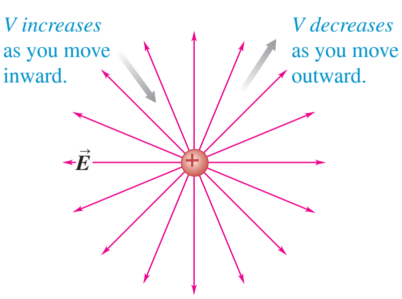
If we divide this by q_0 and compare the result with Eq. (23.13), we find

$$\underbrace{V_a - V_b}_{\text{Electric potential difference}} = \overbrace{\int_a^b \underbrace{\vec{E} \cdot d\vec{l}}_{\substack{\text{Scalar product of electric field}\\ \text{and displacement vector}}}}^{\text{Integral along path from } a \text{ to } b} = \overbrace{\int_a^b \underset{\substack{\text{Electric-field}\\ \text{magnitude}}}{E} \cos \underset{\substack{\text{Angle between}\\ \vec{E} \text{ and } d\vec{l}}}{\phi} \, \overset{\text{Displacement}}{dl}} \qquad (23.17)$$

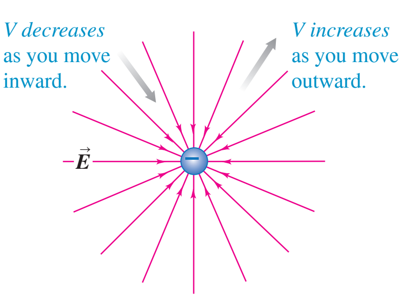
The value of $V_a - V_b$ is independent of the path taken from a to b, just as the value of $W_{a \to b}$ is independent of the path. To interpret Eq. (23.17), remember that $\vec{E}$ is the electric force per unit charge on a test charge. If the line integral $\int_a^b \vec{E} \cdot d\vec{l}$ is positive, the electric field does positive work on a positive test charge as it moves from a to b. In this case the electric potential energy decreases as the test charge moves, so the potential energy per unit charge decreases as well; hence V_b is less than V_a and $V_a - V_b$ is positive.

Figure **23.12** If you move in the direction of $\vec{E}$, electric potential V decreases; if you move in the direction opposite $\vec{E}$, V increases.

(a) A positive point charge



(b) A negative point charge



As an illustration, consider a positive point charge (**Fig. 23.12a**). The electric field is directed away from the charge, and $V = q/4\pi\epsilon_0 r$ is positive at any finite distance from the charge. If you move away from the charge, in the direction of $\vec{E}$, you move toward lower values of V; if you move toward the charge, in the direction opposite $\vec{E}$, you move toward greater values of V. For the negative point charge in Fig. 23.12b, $\vec{E}$ is directed toward the charge and $V = q/4\pi\epsilon_0 r$ is negative at any finite distance from the charge. In this case, if you move toward the charge, you are moving in the direction of $\vec{E}$ and in the direction of decreasing (more negative) V. Moving away from the charge, in the direction opposite $\vec{E}$, moves you toward increasing (less negative) values of V. The general rule, valid for *any* electric field, is: Moving *with* the direction of $\vec{E}$ means moving in the direction of *decreasing* V, and moving *against* the direction of $\vec{E}$ means moving in the direction of *increasing* V.

Also, a positive test charge q_0 experiences an electric force in the direction of $\vec{E}$, toward lower values of V; a negative test charge experiences a force opposite $\vec{E}$, toward higher values of V. Thus a positive charge tends to "fall" from a high-potential region to a lower-potential region. The opposite is true for a negative charge.

Notice that Eq. (23.17) can be rewritten as

$$V_a - V_b = -\int_b^a \vec{E} \cdot d\vec{l} \qquad (23.18)$$

This has a negative sign compared to the integral in Eq. (23.17), and the limits are reversed; hence Eqs. (23.17) and (23.18) are equivalent. But Eq. (23.18) has a slightly different interpretation. To move a unit charge slowly against the electric force, we must apply an *external* force per unit charge equal to $-\vec{E}$, equal and opposite to the electric force per unit charge $\vec{E}$. Equation (23.18) says that $V_a - V_b = V_{ab}$, the potential of a with respect to b, equals the work done per unit charge by this external force to move a unit charge from b to a. This is the same alternative interpretation we discussed under Eq. (23.13).

Equations (23.17) and (23.18) show that the unit of potential difference (1 V) is equal to the unit of electric field (1 N/C) multiplied by the unit of distance (1 m). Hence the unit of electric field can be expressed as 1 *volt per meter* (1 V/m) as well as 1 N/C:

$$1 \text{ V/m} = 1 \text{ volt/meter} = 1 \text{ N/C} = 1 \text{ newton/coulomb}$$

In practice, the volt per meter is the usual unit of electric-field magnitude.

Electron Volts

The magnitude e of the electron charge can be used to define a unit of energy that is useful in many calculations with atomic and nuclear systems. When a particle with charge q moves from a point where the potential is V_b to a point where it is V_a, the change in the potential energy U is

$$U_a - U_b = q(V_a - V_b) = qV_{ab}$$

If charge q equals the magnitude e of the electron charge, 1.602×10^{-19} C, and the potential difference is $V_{ab} = 1$ V $= 1$ J/C, the change in energy is

$$U_a - U_b = (1.602 \times 10^{-19}\ \text{C})(1\ \text{J/C}) = 1.602 \times 10^{-19}\ \text{J}$$

This quantity of energy is defined to be 1 **electron volt** (1 eV):

$$1\ \text{eV} = 1.602 \times 10^{-19}\ \text{J}$$

The multiples meV, keV, MeV, GeV, and TeV are often used.

CAUTION Electron volts vs. volts Remember that the electron volt is a unit of energy, *not* a unit of potential or potential difference!

When a particle with charge e moves through a potential difference of 1 volt, the change in potential *energy* is 1 eV. If the charge is some multiple of e—say, Ne—the change in potential energy in electron volts is N times the potential difference in volts. For example, when an alpha particle, which has charge $2e$, moves between two points with a potential difference of 1000 V, the change in potential energy is $2(1000\ \text{eV}) = 2000$ eV. To confirm this, we write

$$\begin{aligned} U_a - U_b = qV_{ab} &= (2e)(1000\ \text{V}) = (2)(1.602 \times 10^{-19}\ \text{C})(1000\ \text{V}) \\ &= 3.204 \times 10^{-16}\ \text{J} = 2000\ \text{eV} \end{aligned}$$

Although we defined the electron volt in terms of *potential* energy, we can use it for *any* form of energy, such as the kinetic energy of a moving particle. When we speak of a "one-million-electron-volt proton," we mean a proton with a kinetic energy of one million electron volts (1 MeV), equal to $(10^6)(1.602 \times 10^{-19}\ \text{J}) = 1.602 \times 10^{-13}$ J. The Large Hadron Collider near Geneva, Switzerland, is designed to accelerate protons to a kinetic energy of 7 TeV (7×10^{12} eV).

BIO APPLICATION Electron Volts and Cancer Radiotherapy One way to destroy a cancerous tumor is to aim high-energy electrons directly at it. Each electron has a kinetic energy of 4 to 20 million electron volts, or MeV (1 MeV $= 10^6$ eV), and transfers its energy to the tumor through collisions with the tumor's atoms. Electrons in this energy range can penetrate only a few centimeters into a patient, which makes them useful for treating superficial tumors, such as those on the skin or lips.

EXAMPLE 23.3 Electric force and electric potential

WITH VARIATION PROBLEMS

A proton (charge $+e = 1.602 \times 10^{-19}$ C) moves a distance $d = 0.50$ m in a straight line between points a and b in a linear accelerator. The electric field is uniform along this line, with magnitude $E = 1.5 \times 10^7$ V/m $= 1.5 \times 10^7$ N/C in the direction from a to b. Determine (a) the force on the proton; (b) the work done on it by the field; (c) the potential difference $V_a - V_b$.

IDENTIFY and SET UP This problem uses the relationship between electric field and electric force. It also uses the relationship among force, work, and potential-energy difference. We are given the electric field, so it is straightforward to find the electric force on the proton. Calculating the work is also straightforward because $\vec{E}$ is uniform, so the force on the proton is constant. Once the work is known, we find $V_a - V_b$ from Eq. (23.13).

EXECUTE (a) The force on the proton is in the same direction as the electric field, and its magnitude is

$$\begin{aligned} F = qE &= (1.602 \times 10^{-19}\ \text{C})(1.5 \times 10^7\ \text{N/C}) \\ &= 2.4 \times 10^{-12}\ \text{N} \end{aligned}$$

(b) The force is constant and in the same direction as the displacement, so the work done on the proton is

$$\begin{aligned} W_{a\to b} = Fd &= (2.4 \times 10^{-12}\ \text{N})(0.50\ \text{m}) \\ &= 1.2 \times 10^{-12}\ \text{J} \\ &= (1.2 \times 10^{-12}\ \text{J})\frac{1\ \text{eV}}{1.602 \times 10^{-19}\ \text{J}} \\ &= 7.5 \times 10^6\ \text{eV} = 7.5\ \text{MeV} \end{aligned}$$

(c) From Eq. (23.13) the potential difference is the work per unit charge, which is

$$\begin{aligned} V_a - V_b &= \frac{W_{a\to b}}{q} = \frac{1.2 \times 10^{-12}\ \text{J}}{1.602 \times 10^{-19}\ \text{C}} \\ &= 7.5 \times 10^6\ \text{J/C} = 7.5 \times 10^6\ \text{V} = 7.5\ \text{MV} \end{aligned}$$

We can get this same result even more easily by remembering that 1 electron volt equals 1 volt multiplied by the charge e. The work done is 7.5×10^6 eV and the charge is e, so the potential difference is $(7.5 \times 10^6\ \text{eV})/e = 7.5 \times 10^6$ V.

EVALUATE We can check our result in part (c) by using Eq. (23.17) or (23.18). The angle ϕ between the constant field $\vec{E}$ and the displacement is zero, so Eq. (23.17) becomes

$$V_a - V_b = \int_a^b E\cos\phi\, dl = \int_a^b E\, dl = E\int_a^b dl$$

The integral of dl from a to b is just the distance d, so we again find

$$V_a - V_b = Ed = (1.5 \times 10^7\ \text{V/m})(0.50\ \text{m}) = 7.5 \times 10^6\ \text{V}$$

KEYCONCEPT The potential difference V_{ab} between point a and point b, equal to the difference $V_a - V_b$ of the potentials V at the two points, is the amount of work the electric force does on a unit charge as it moves from a to b. If you move in the direction of the electric field, V decreases (V_{ab} is positive); if you move opposite to the direction of the electric field, V increases (V_{ab} is negative).

EXAMPLE 23.4 Potential due to two point charges

WITH VARIATION PROBLEMS

An electric dipole consists of point charges $q_1 = +12\text{ nC}$ and $q_2 = -12\text{ nC}$ placed 10.0 cm apart (**Fig. 23.13**). Compute the electric potentials at points a, b, and c.

IDENTIFY and SET UP This is the same arrangement as in Example 21.8 (Section 21.5), in which we calculated the electric *field* at each point by doing a *vector* sum. Here our target variable is the electric *potential* V at three points, which we find by doing the *algebraic* sum in Eq. (23.15).

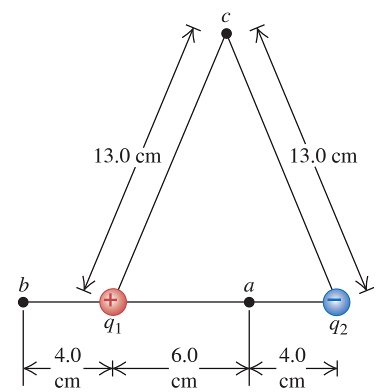
EXECUTE At point a we have $r_1 = 0.060\text{ m}$ and $r_2 = 0.040\text{ m}$, so Eq. (23.15) becomes

$$\begin{aligned} V_a &= \frac{1}{4\pi\epsilon_0}\sum_i \frac{q_i}{r_i} = \frac{1}{4\pi\epsilon_0}\frac{q_1}{r_1} + \frac{1}{4\pi\epsilon_0}\frac{q_2}{r_2} \\ &= (9.0 \times 10^9\text{ N}\cdot\text{m}^2/\text{C}^2)\frac{12 \times 10^{-9}\text{ C}}{0.060\text{ m}} \\ &\quad + (9.0 \times 10^9\text{ N}\cdot\text{m}^2/\text{C}^2)\frac{(-12 \times 10^{-9}\text{ C})}{0.040\text{ m}} \\ &= 1800\text{ N}\cdot\text{m/C} + (-2700\text{ N}\cdot\text{m/C}) \\ &= 1800\text{ V} + (-2700\text{ V}) = -900\text{ V} \end{aligned}$$

In a similar way you can show that the potential at point b (where $r_1 = 0.040\text{ m}$ and $r_2 = 0.140\text{ m}$) is $V_b = 1930\text{ V}$ and that the potential at point c (where $r_1 = r_2 = 0.130\text{ m}$) is $V_c = 0$.

EVALUATE Let's confirm that these results make sense. Point a is closer to the -12 nC charge than to the $+12$ nC charge, so the potential at a is negative. The potential is positive at point b, which is closer to the $+12$ nC charge than the -12 nC charge. Finally, point c is equidistant from the $+12$ nC charge and the -12 nC charge, so the potential there is zero. (The potential is also equal to zero at a point infinitely far from both charges.)

Figure 23.13 What are the potentials at points a, b, and c due to this electric dipole?



Comparing this example with Example 21.8 shows that it's much easier to calculate electric potential (a scalar) than electric field (a vector). We'll take advantage of this simplification whenever possible.

KEYCONCEPT The electric potential caused at a point P by a single point charge is proportional to the charge and inversely proportional to the distance from the point charge to P. The total electric potential at P due to a system of charges is the sum of the potentials at P due to each individual charge.

EXAMPLE 23.5 Potential and potential energy

WITH VARIATION PROBLEMS

Compute the potential energy associated with a $+4.0$ nC point charge if it is placed at points a, b, and c in Fig. 23.13.

IDENTIFY and SET UP The potential energy U associated with a point charge q at a location where the electric potential is V is $U = qV$. We use the values of V from Example 23.4.

EXECUTE At the three points we find

$$U_a = qV_a = (4.0 \times 10^{-9}\text{ C})(-900\text{ J/C}) = -3.6 \times 10^{-6}\text{ J}$$

$$U_b = qV_b = (4.0 \times 10^{-9}\text{ C})(1930\text{ J/C}) = 7.7 \times 10^{-6}\text{ J}$$

$$U_c = qV_c = 0$$

All of these values correspond to U and V being zero at infinity.

EVALUATE Note that *zero* net work is done on the 4.0 nC charge if it moves from point c to infinity *by any path*. In particular, let the path be along the perpendicular bisector of the line joining the other two charges q_1 and q_2 in Fig. 23.13. As shown in Example 21.8 (Section 21.5), at points on the bisector, the direction of $\vec{\boldsymbol{E}}$ is perpendicular to the bisector. Hence the force on the 4.0 nC charge is perpendicular to the path, and no work is done in any displacement along it.

KEYCONCEPT If you place a point charge q at a point P, the associated electric potential energy is equal to q multiplied by the value at P of the electric potential due to all *other* charges.

EXAMPLE 23.6 Finding potential by integration

By integrating the electric field as in Eq. (23.17), find the potential at a distance r from a point charge q.

IDENTIFY and SET UP We let point a in Eq. (23.17) be at distance r and let point b be at infinity (**Fig. 23.14**). As usual, we choose the potential to be zero at an infinite distance from the charge q.

EXECUTE To carry out the integral, we can choose any path we like between points a and b. The most convenient path is a radial line as shown in Fig. 23.14, so that $d\vec{l}$ is in the radial direction and has magnitude dr. Writing $d\vec{l} = \hat{r}dr$, we have from Eq. (23.17)

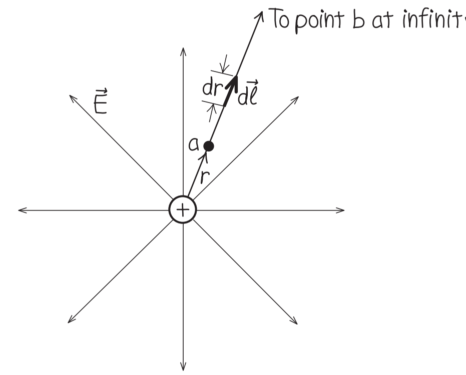
$$V - 0 = V = \int_r^\infty \vec{E} \cdot d\vec{l}$$

$$= \int_r^\infty \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \hat{r} \cdot \hat{r} dr = \int_r^\infty \frac{q}{4\pi\epsilon_0 r^2} dr$$

$$= -\frac{q}{4\pi\epsilon_0 r}\bigg|_r^\infty = 0 - \left(-\frac{q}{4\pi\epsilon_0 r}\right) = \frac{q}{4\pi\epsilon_0 r}$$

EVALUATE Our result agrees with Eq. (23.14) and is correct for positive or negative q.

Figure **23.14** Calculating the potential by integrating $\vec{E}$ for a single point charge.



KEYCONCEPT The potential difference $V_{ab} = V_a - V_b$ between point a and point b equals the integral of the electric field along a path from a to b. This integral does not depend on the path taken between the two points.

EXAMPLE 23.7 Moving through a potential difference

WITH VARIATION PROBLEMS

In **Fig. 23.15** a dust particle with mass $m = 5.0 \times 10^{-9}$ kg $= 5.0\ \mu$g and charge $q_0 = 2.0$ nC starts from rest and moves in a straight line from point a to point b. What is its speed v at point b?

IDENTIFY and SET UP Only the conservative electric force acts on the particle, so the total mechanical energy is conserved: $K_a + U_a = K_b + U_b$. We get the potential energies U from the corresponding potentials V from Eq. (23.12): $U_a = q_0 V_a$ and $U_b = q_0 V_b$.

EXECUTE We have $K_a = 0$ and $K_b = \frac{1}{2}mv^2$. We substitute these and our expressions for U_a and U_b into the energy-conservation equation, then solve for v. We find

$$0 + q_0 V_a = \tfrac{1}{2}mv^2 + q_0 V_b$$

$$v = \sqrt{\frac{2q_0(V_a - V_b)}{m}}$$

We calculate the potentials from Eq. (23.15), $V = q/4\pi\epsilon_0 r$:

$$V_a = (9.0 \times 10^9\ \mathrm{N \cdot m^2/C^2})\left(\frac{3.0 \times 10^{-9}\ \mathrm{C}}{0.010\ \mathrm{m}} + \frac{(-3.0 \times 10^{-9}\ \mathrm{C})}{0.020\ \mathrm{m}}\right)$$

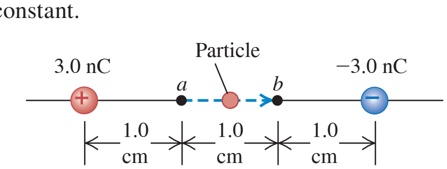
$$= 1350\ \mathrm{V}$$

$$V_b = (9.0 \times 10^9\ \mathrm{N \cdot m^2/C^2})\left(\frac{3.0 \times 10^{-9}\ \mathrm{C}}{0.020\ \mathrm{m}} + \frac{(-3.0 \times 10^{-9}\ \mathrm{C})}{0.010\ \mathrm{m}}\right)$$

$$= -1350\ \mathrm{V}$$

$$V_a - V_b = (1350\ \mathrm{V}) - (-1350\ \mathrm{V}) = 2700\ \mathrm{V}$$

Figure **23.15** The particle moves from point a to point b; its acceleration is not constant.



Finally,

$$v = \sqrt{\frac{2(2.0 \times 10^{-9}\ \mathrm{C})(2700\ \mathrm{V})}{5.0 \times 10^{-9}\ \mathrm{kg}}} = 46\ \mathrm{m/s}$$

EVALUATE Our result makes sense: The positive dust particle speeds up as it moves away from the +3.0 nC charge and toward the −3.0 nC charge. To check unit consistency in the final line of the calculation, note that 1 V = 1 J/C, so the numerator under the radical has units of J or kg · m^2/s^2.

KEYCONCEPT When a particle of charge q moves from a point of higher potential to a point of lower potential, its potential energy *decreases* if q is positive; its potential energy *increases* if q is negative.

TEST YOUR UNDERSTANDING OF SECTION 23.2 If the electric *potential* at a certain point is zero, does the electric *field* at that point have to be zero? (*Hint:* Consider point c in Examples 23.4 and 21.8.)

ANSWER

no If $V = 0$ at a certain point, $\vec{E}$ does *not* have to be zero at that point. An example is point c in Figs. 21.23 and 23.13, for which there is an electric field in the $+x$-direction (see Example 21.9 in Section 21.5) even though $V = 0$ (see Example 23.4). This isn't a surprising result because V and $\vec{E}$ are quite different quantities: V is the net amount of work required to bring a unit charge from infinity to the point in question, whereas $\vec{E}$ is the electric force that acts on a unit charge when it arrives at that point.

23.3 CALCULATING ELECTRIC POTENTIAL

When calculating the potential due to a charge distribution, we usually follow one of two routes. If we know the charge distribution, we can use Eq. (23.15) or (23.16). Or if we know how the electric field depends on position, we can use Eq. (23.17), defining the potential to be zero at some convenient place. Some problems require a combination of these approaches.

As you read through these examples, compare them with the related examples of calculating electric *field* in Section 21.5. You'll see how much easier it is to calculate scalar electric potentials than vector electric fields. The moral is clear: Whenever possible, solve problems by means of an energy approach (using electric potential and electric potential energy) rather than a dynamics approach (using electric fields and electric forces).

PROBLEM-SOLVING STRATEGY 23.1 Calculating Electric Potential

IDENTIFY *the relevant concepts:* Remember that electric potential is *potential energy per unit charge*.

SET UP *the problem* using the following steps:

1. Make a drawing showing the locations and values of the charges (which may be point charges or a continuous distribution of charge) and your choice of coordinate axes.
2. Indicate on your drawing the position of the point at which you want to calculate the electric potential V. Sometimes this position will be an arbitrary one (say, a point a distance r from the center of a charged sphere).

EXECUTE *the solution* as follows:

1. To find the potential due to a collection of point charges, use Eq. (23.15). If you are given a continuous charge distribution, devise a way to divide it into infinitesimal elements and use Eq. (23.16). Carry out the integration, using appropriate limits to include the entire charge distribution.
2. If you are given the electric field, or if you can find it from any of the methods presented in Chapters 21 or 22, it may be easier to find the potential difference between points a and b from Eq. (23.17) or (23.18). When appropriate, make use of your freedom to define V to be zero at some convenient place, and choose this place to be point b. (For point charges, this will usually be at infinity. For other distributions of charge—especially those that themselves extend to infinity—it may be necessary to define V_b to be zero at some finite distance from the charge distribution.) Then the potential at any other point, say a, can by found from Eq. (23.17) or (23.18) with $V_b = 0$.
3. Although potential V is a *scalar* quantity, you may have to use components of the vectors $\vec{E}$ and $d\vec{l}$ when you use Eq. (23.17) or (23.18) to calculate V.

EVALUATE *your answer:* Check whether your answer agrees with your intuition. If your result gives V as a function of position, graph the function to see whether it makes sense. If you know the electric field, you can make a rough check of your result for V by verifying that V decreases if you move in the direction of $\vec{E}$.

EXAMPLE 23.8 A charged conducting sphere

WITH VARIATION PROBLEMS

A solid conducting sphere of radius R has a total charge q. Find the electric potential everywhere, both outside and inside the sphere.

IDENTIFY and SET UP In Example 22.5 (Section 22.4) we used Gauss's law to find the electric *field* at all points for this charge distribution. We can use that result to determine the potential.

EXECUTE From Example 22.5, the field *outside* the sphere is the same as if the sphere were removed and replaced by a point charge q. We take $V = 0$ at infinity, as we did for a point charge. Then the potential at a point outside the sphere at a distance r from its center is the same as that due to a point charge q at the center:

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{r}$$

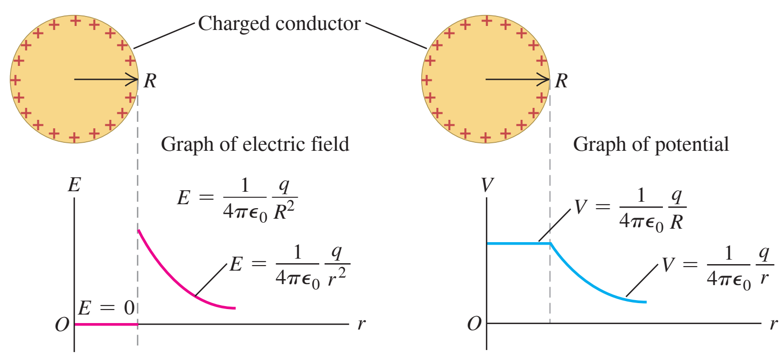
The potential at the surface of the sphere is $V_{\text{surface}} = q/4\pi\epsilon_0 R$.

Inside the sphere, $\vec{E}$ is zero everywhere. Hence no work is done on a test charge that moves from any point to any other point inside the sphere. Thus the potential is the same at every point inside the sphere and is equal to its value $q/4\pi\epsilon_0 R$ at the surface.

EVALUATE **Figure 23.16** shows the field and potential for a positive charge q. In this case the electric field points radially away from the sphere. As you move away from the sphere, in the direction of $\vec{E}$, V decreases (as it should).

KEYCONCEPT If there is excess charge at rest on a solid spherical conductor, all of the charge lies on the surface of the conductor, and the electric field and electric potential outside the sphere are the same as if all the charge were concentrated at the center of the sphere. Everywhere inside the sphere the field is zero and the potential has the same value as at the surface.

Figure 23.16 Electric-field magnitude E and potential V at points inside and outside a positively charged spherical conductor.



Ionization and Corona Discharge

The results of Example 23.8 have numerous practical consequences. One consequence relates to the maximum potential to which a conductor in air can be raised. This potential is limited because air molecules become *ionized,* and air becomes a conductor, at an electric-field magnitude of about 3×10^6 V/m. Assume for the moment that q is positive. When we compare the expressions in Example 23.8 for the potential V_{surface} and field magnitude E_{surface} at the surface of a charged conducting sphere, we note that $V_{\text{surface}} = E_{\text{surface}}R$. Thus, if E_{m} represents the electric-field magnitude at which air becomes conductive (known as the *dielectric strength* of air), then the maximum potential V_{m} to which a spherical conductor can be raised is

$$V_{\text{m}} = RE_{\text{m}}$$

For a conducting sphere 1 cm in radius in air, $V_{\text{m}} = (10^{-2}\text{ m})(3 \times 10^6\text{ V/m}) = 30{,}000$ V. No amount of "charging" could raise the potential of a conducting sphere of this size in air higher than about 30,000 V; attempting to raise the potential further by adding extra charge would cause the surrounding air to become ionized and conductive, and the extra added charge would leak into the air.

To attain even higher potentials, high-voltage machines such as Van de Graaff generators use spherical terminals with very large radii (see Fig. 22.26 and the photograph that opens Chapter 22). For example, a terminal of radius $R = 2$ m has a maximum potential $V_{\text{m}} = (2\text{ m})(3 \times 10^6\text{ V/m}) = 6 \times 10^6\text{ V} = 6$ MV.

Figure 23.17 The metal mast at the top of the Empire State Building acts as a lightning rod. It is struck by lightning as many as 500 times each year.

Our result in Example 23.8 also explains what happens with a charged conductor with a very *small* radius of curvature, such as a sharp point or thin wire. Because the maximum potential is proportional to the radius, even relatively small potentials applied to sharp points in air produce sufficiently high fields just outside the point to ionize the surrounding air, making it become a conductor. The resulting current and its associated glow (visible in a dark room) are called *corona discharge*. Laser printers and photocopying machines use corona discharge from fine wires to spray charge on the imaging drum (see Fig. 21.2).

A large-radius conductor is used in situations where it's important to *prevent* corona discharge. An example is the blunt end of a metal lightning rod (**Fig. 23.17**). If there is an excess charge in the atmosphere, as happens during thunderstorms, a substantial charge of the opposite sign can build up on this blunt end. As a result, when the atmospheric charge is discharged through a lightning bolt, it tends to be attracted to the charged lightning rod rather than to other structures that could be damaged. (A conducting wire connecting the lightning rod to the ground then allows the acquired charge to dissipate harmlessly.) A lightning rod with a sharp end would allow less charge buildup and hence be less effective.

EXAMPLE 23.9 Oppositely charged parallel plates

WITH VARIATION PROBLEMS

Find the potential at any height y between the two oppositely charged parallel plates discussed in Section 23.1 (**Fig. 23.18**).

IDENTIFY and SET UP We discussed this situation in Section 23.1. From Eq. (23.5), we know the electric *potential energy* U for a test charge q_0 is $U = q_0Ey$. (We set $y = 0$ and $U = 0$ at the bottom plate.) We use Eq. (23.12), $U = q_0V$, to find the electric *potential* V as a function of y.

EXECUTE The potential $V(y)$ at coordinate y is the potential energy per unit charge:

$$V(y) = \frac{U(y)}{q_0} = \frac{q_0Ey}{q_0} = Ey$$

The potential decreases as we move in the direction of $\vec{E}$ from the upper to the lower plate. At point a, where $y = d$ and $V(y) = V_a$,

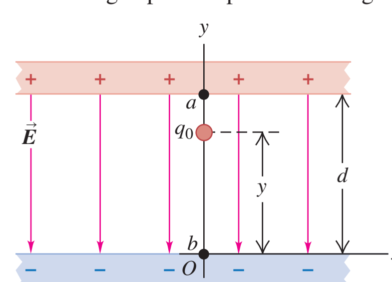
$$V_a - V_b = Ed$$

and

$$E = \frac{V_a - V_b}{d} = \frac{V_{ab}}{d}$$

where V_{ab} is the potential of the positive plate with respect to the negative plate. That is, the electric field equals the potential difference between the plates divided by the distance between them. For a given potential difference V_{ab}, the smaller the distance d between the two plates, the greater the magnitude E of the electric field. (This relationship between E and V_{ab} holds *only* for the planar geometry we have described. It does *not* work for situations such as concentric cylinders or spheres in which the electric field is not uniform.) ?

Figure 23.18 The charged parallel plates from Fig. 23.2.



EVALUATE Our result shows that $V = 0$ at the bottom plate (at $y = 0$). This is consistent with our choice that $U = q_0V = 0$ for a test charge placed at the bottom plate.

CAUTION "Zero potential" is arbitrary You might think that if a conducting object has zero potential, it must also have zero net charge. But that just isn't so! As an example, the plate at $y = 0$ in Fig. 23.18 has zero potential $(V = 0)$ but has a nonzero charge per unit area $-\sigma$. There's nothing special about the place where potential is zero; we *define* this place to be wherever we want it to be.

KEYCONCEPT In a region of uniform electric field $\vec{E}$, such as between two oppositely charged parallel plates, the electric potential decreases linearly with position as you move in the direction of $\vec{E}$.

EXAMPLE 23.10 An infinite line charge or charged conducting cylinder

WITH VARIATION PROBLEMS

Find the potential at a distance r from a very long line of charge with linear charge density (charge per unit length) λ.

IDENTIFY and SET UP In both Example 21.10 (Section 21.5) and Example 22.6 (Section 22.4) we found that the electric field at a radial distance r from a long straight-line charge (**Fig. 23.19a**) has only a radial component $E_r = \lambda/2\pi\epsilon_0 r$. We use this expression to find the potential by integrating $\vec{E}$ as in Eq. (23.17).

EXECUTE Since the field has only a radial component, we have $\vec{E}\cdot d\vec{l} = E_r dr$. Hence from Eq. (23.17) the potential of any point a with respect to any other point b, at radial distances r_a and r_b from the line of charge, is

$$V_a - V_b = \int_a^b \vec{E}\cdot d\vec{l} = \int_a^b E_r dr = \frac{\lambda}{2\pi\epsilon_0}\int_{r_a}^{r_b}\frac{dr}{r} = \frac{\lambda}{2\pi\epsilon_0}\ln\frac{r_b}{r_a}$$

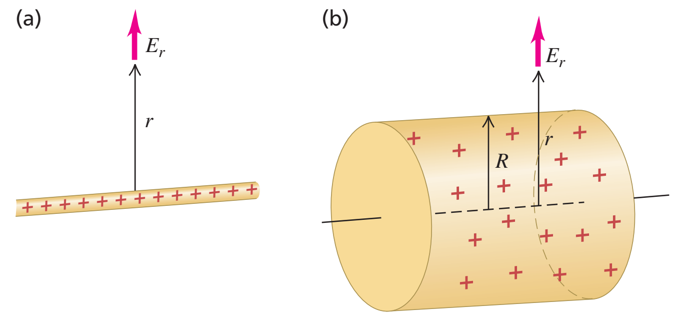
If we take point b at infinity and set $V_b = 0$, we find that V_a is *infinite* for any finite distance r_a from the line charge: $V_a = (\lambda/2\pi\epsilon_0)\ln(\infty/r_a) = \infty$. This is *not* a useful way to define V for this problem! The difficulty is that the charge distribution itself extends to infinity.

Instead, as recommended in Problem-Solving Strategy 23.1, we set $V_b = 0$ at point b at an arbitrary but *finite* radial distance r_0. Then the potential $V = V_a$ at point a at a radial distance r is given by $V - 0 = (\lambda/2\pi\epsilon_0)\ln(r_0/r)$, or

$$V = \frac{\lambda}{2\pi\epsilon_0}\ln\frac{r_0}{r}$$

EVALUATE According to our result, if λ is positive, then V decreases as r increases. This is as it should be: V decreases as we move in the direction of $\vec{E}$.

Figure **23.19** Electric field outside (a) a long, positively charged wire and (b) a long, positively charged cylinder.



From Example 22.6, the expression for E_r with which we started also applies outside a long, charged conducting cylinder with charge per unit length λ (Fig. 23.19b). Hence our result also gives the potential for such a cylinder, but only for values of r (the distance from the cylinder axis) equal to or greater than the radius R of the cylinder. If we choose r_0 to be the radius R, so that $V = 0$ when $r = R$, then at any point for which $r > R$,

$$V = \frac{\lambda}{2\pi\epsilon_0}\ln\frac{R}{r}$$

Inside the cylinder, $\vec{E} = \mathbf{0}$, and V has the same value (zero) as on the cylinder's surface.

KEYCONCEPT The electric potential outside a very long cylindrical distribution of charge or line of charge decreases logarithmically with the distance from the axis of the distribution or line.

EXAMPLE 23.11 A ring of charge

WITH VARIATION PROBLEMS

Electric charge Q is distributed uniformly around a thin ring of radius a (**Fig. 23.20**). Find the potential at a point P on the ring axis at a distance x from the center of the ring.

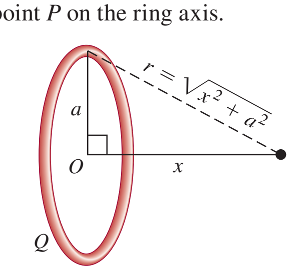
IDENTIFY and SET UP We divide the ring into infinitesimal segments and use Eq. (23.16) to find V. All parts of the ring (and therefore all elements of the charge distribution) are at the same distance from P.

EXECUTE Figure 23.20 shows that the distance from each charge element dq to P is $r = \sqrt{x^2 + a^2}$. Hence we can take the factor $1/r$ outside the integral in Eq. (23.16), and

$$\begin{aligned} V &= \frac{1}{4\pi\epsilon_0}\int\frac{dq}{r} \\ &= \frac{1}{4\pi\epsilon_0}\frac{1}{\sqrt{x^2+a^2}}\int dq \\ &= \frac{1}{4\pi\epsilon_0}\frac{Q}{\sqrt{x^2+a^2}} \end{aligned}$$

EVALUATE When x is much larger than a, our expression for V becomes approximately $V = Q/4\pi\epsilon_0 x$, which is the potential at a distance x from a point charge Q. Very far from a charged ring, its electric potential looks like that of a point charge. We drew a similar conclusion about the electric *field* of a ring in Example 21.9 (Section 21.5).

Figure **23.20** All the charge in a ring of charge Q is the same distance r from a point P on the ring axis.



We know the electric field at all points along the x-axis from Example 21.9 (Section 21.5), so we can also find V along this axis by integrating $\vec{E}\cdot d\vec{l}$ as in Eq. (23.17).

KEYCONCEPT To find the electric potential at a point due to a continuous distribution of charge, first divide the distribution into infinitesimally small segments. Then find the potential at the point due to one such segment. Finally, integrate over all segments in the charge distribution.

EXAMPLE 23.12 Potential of a line of charge

WITH VARIATION PROBLEMS

Positive electric charge Q is distributed uniformly along a line of length $2a$ lying along the y-axis between $y = -a$ and $y = +a$ (**Fig. 23.21**). Find the electric potential at a point P on the x-axis at a distance x from the origin.

IDENTIFY and SET UP This is the situation of Example 21.10 (Section 21.5), where we found an expression for the electric field $\vec{E}$ at an arbitrary point on the x-axis. We can find V at point P by using Eq. (23.16) to integrate over the charge distribution. Unlike the situation in Example 23.11, each charge element dQ is a *different* distance from point P, so the integration will take a little more effort.

EXECUTE As in Example 21.10, the element of charge dQ corresponding to an element of length dy on the rod is $dQ = (Q/2a)dy$. The distance from dQ to P is $\sqrt{x^2 + y^2}$, so the contribution dV that the charge element makes to the potential at P is

$$dV = \frac{1}{4\pi\epsilon_0}\frac{Q}{2a}\frac{dy}{\sqrt{x^2 + y^2}}$$

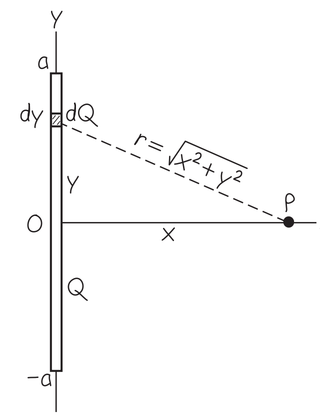
To find the potential at P due to the entire rod, we integrate dV over the length of the rod from $y = -a$ to $y = a$:

$$V = \frac{1}{4\pi\epsilon_0}\frac{Q}{2a}\int_{-a}^{a}\frac{dy}{\sqrt{x^2 + y^2}}$$

You can look up the integral in a table. The final result is

$$V = \frac{1}{4\pi\epsilon_0}\frac{Q}{2a}\ln\left(\frac{\sqrt{a^2 + x^2} + a}{\sqrt{a^2 + x^2} - a}\right)$$

Figure 23.21 Our sketch for this problem.



EVALUATE We can check our result by letting x approach infinity. In this limit the point P is infinitely far from all of the charge, so we expect V to approach zero; you can verify that it does.

We know the electric field at all points along the x-axis from Example 21.10. We invite you to use this information to find V along this axis by integrating $\vec{E}$ as in Eq. (23.17).

KEYCONCEPT The electric potential due to a symmetrical distribution of charge is most easily calculated at a point of symmetry. Whenever possible, take advantage of the symmetry of the situation to check your results.

TEST YOUR UNDERSTANDING OF SECTION 23.3 If the electric *field* at a certain point is zero, does the electric *potential* at that point have to be zero? (*Hint:* Consider the center of the ring in Examples 23.11 and 21.9.)

ANSWER

no If $\vec{E} = \mathbf{0}$ at a certain point, V does *not* have to be zero at that point. An example is point O at the center of the charged ring in Figs. 21.23 and 23.21. From Example 21.9 (Section 21.5), the electric field is zero at O because the electric-field contributions from different parts of the ring completely cancel. From Example 23.11, however, the potential at O is *not* zero: This point corresponds to $x = 0$, so $V = (1/4\pi\epsilon_0)(Q/a)$. This value of V corresponds to the work that would have to be done to move a unit positive test charge along a path from infinity to point O; it is nonzero because the charged ring repels the test charge, so positive work must be done to move the test charge toward the ring.

23.4 EQUIPOTENTIAL SURFACES

Field lines (see Section 21.6) help us visualize electric fields. In a similar way, the potential at various points in an electric field can be represented graphically by *equipotential surfaces*. These use the same fundamental idea as topographic maps like those used by hikers and mountain climbers (**Fig. 23.22**). On a topographic map, contour lines are drawn through points that are all at the same elevation. Any number of these could be drawn, but typically only a few contour lines are shown at equal spacings of elevation. If a mass m is moved over the terrain along such a contour line, the gravitational potential energy mgy does not change because the elevation y is constant. Thus contour lines on a topographic map are really curves of constant gravitational potential energy. Contour lines are close together where the terrain is steep and there are large changes in elevation over a small horizontal distance; the contour lines are farther apart where the terrain

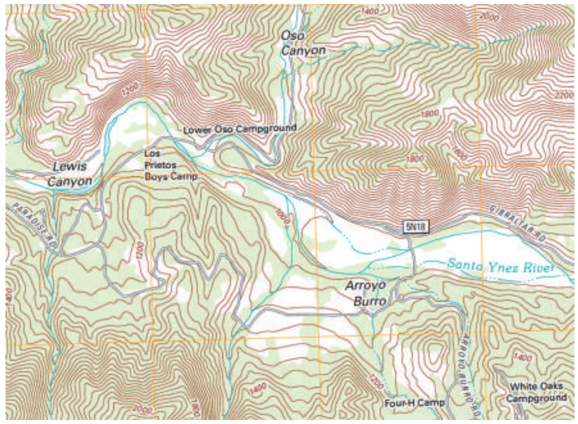


Figure **23.22** Contour lines on a topographic map are curves of constant elevation and hence of constant gravitational potential energy.

is gently sloping. A ball allowed to roll downhill will experience the greatest downhill gravitational force where contour lines are closest together.

By analogy to contour lines on a topographic map, an **equipotential surface** is a three-dimensional surface on which the *electric potential V* is the same at every point. If a test charge q_0 is moved from point to point on such a surface, the *electric* potential energy $q_0 V$ remains constant. In a region where an electric field is present, we can construct an equipotential surface through any point. In diagrams we usually show only a few representative equipotentials, often with equal potential differences between adjacent surfaces. No point can be at two different potentials, so equipotential surfaces for different potentials can never touch or intersect.

Equipotential Surfaces and Field Lines

Because potential energy does not change as a test charge moves over an equipotential surface, the electric field can do no work on such a charge. It follows that $\vec{E}$ must be perpendicular to the surface at every point so that the electric force $q_0\vec{E}$ is always perpendicular to the displacement of a charge moving on the surface. **Field lines and equipotential surfaces are always mutually perpendicular.** In general, field lines are curves, and equipotentials are curved surfaces. For the special case of a *uniform* field, in which the field lines are straight, parallel, and equally spaced, the equipotentials are parallel *planes* perpendicular to the field lines.

Figure 23.23 shows three arrangements of charges. The field lines in the plane of the charges are represented by red lines, and the intersections of the equipotential surfaces with this plane (that is, cross sections of these surfaces) are shown as blue lines. The actual equipotential surfaces are three-dimensional. At each crossing of an equipotential and a field line, the two are perpendicular.

Figure **23.23** Cross sections of equipotential surfaces (blue lines) and electric field lines (red lines) for assemblies of point charges. There are equal potential differences between adjacent surfaces. Compare these diagrams to those in Fig. 21.28, which showed only the electric field lines.

(a) A single positive charge (b) An electric dipole (c) Two equal positive charges

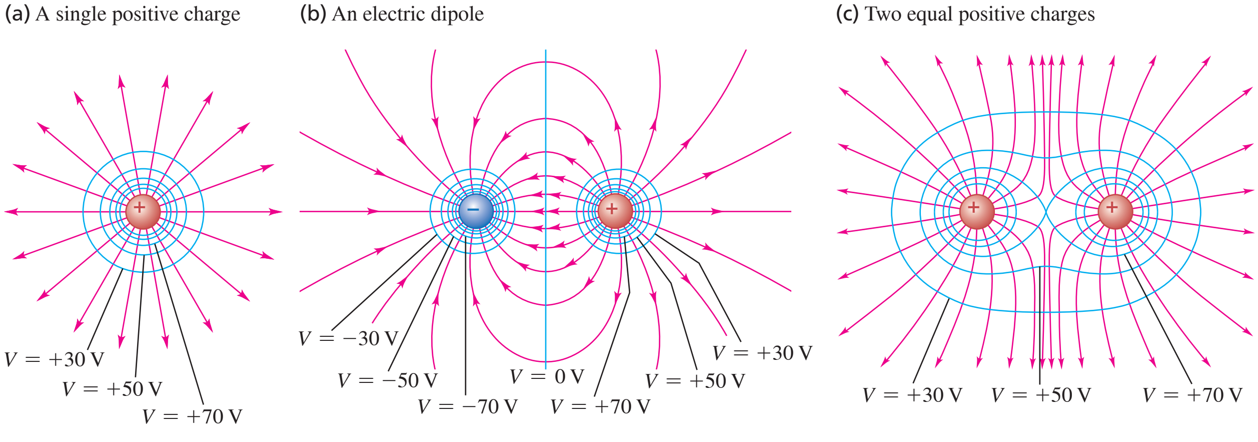


Figure **23.24** When charges are at rest, a conducting surface is always an equipotential surface. Field lines are perpendicular to a conducting surface.

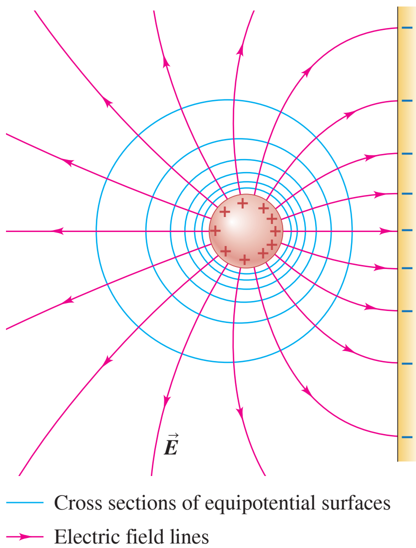


Figure **23.25** At all points on a conductor's surface, the electric field must be perpendicular to the surface. If $\vec{E}$ had a tangential component, a net amount of work would be done on a test charge by moving it around a loop as shown—which is impossible because the electric force is conservative.

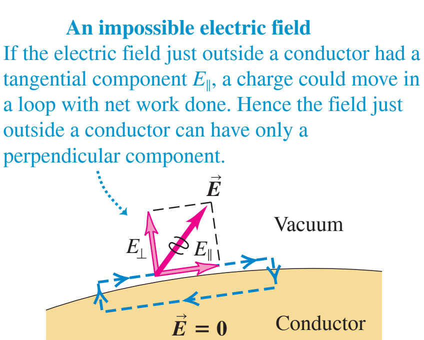
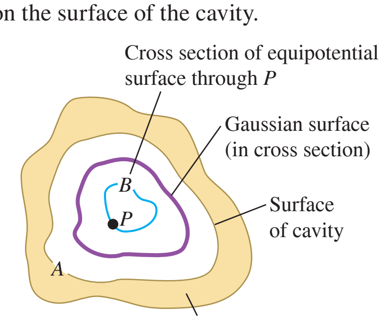


Figure **23.26** A cavity in a conductor. If the cavity contains no charge, every point in the cavity is at the same potential, the electric field is zero everywhere in the cavity, and there is no charge anywhere on the surface of the cavity.



In Fig. 23.23 we have drawn equipotentials so that there are equal potential differences between adjacent surfaces. In regions where the magnitude of $\vec{E}$ is large, the equipotential surfaces are close together because the field does a relatively large amount of work on a test charge in a relatively small displacement. This is the case near the point charge in Fig. 23.23a or between the two point charges in Fig. 23.23b; note that in these regions the field lines are also closer together. This is directly analogous to the downhill force of gravity being greatest in regions on a topographic map where contour lines are close together. Conversely, in regions where the field is weaker, the equipotential surfaces are farther apart; this happens at larger radii in Fig. 23.23a, to the left of the negative charge or the right of the positive charge in Fig. 23.23b, and at greater distances from both charges in Fig. 23.23c. (It may appear that two equipotential surfaces intersect at the center of Fig. 23.23c, in violation of the rule that this can never happen. In fact this is a single figure-8–shaped equipotential surface.)

CAUTION ***E* need not be constant over an equipotential surface** On a given equipotential surface, the potential V has the same value at every point. In general, however, the electric-field magnitude E is *not* the same at all points on an equipotential surface. For instance, on equipotential surface "$V = -30$ V" in Fig. 23.23b, E is less to the left of the negative charge than it is between the two charges. On the figure-8–shaped equipotential surface in Fig. 23.23c, $E = 0$ at the middle point halfway between the two charges; at any other point on this surface, E is nonzero.

Equipotentials and Conductors

Here's an important statement about equipotential surfaces: **When all charges are at rest, the surface of a conductor is always an equipotential surface.** Since the electric field $\vec{E}$ is always perpendicular to an equipotential surface, we can prove this statement by proving that **when all charges are at rest, the electric field just outside a conductor must be perpendicular to the surface at every point** (**Fig. 23.24**). We know that $\vec{E} = \mathbf{0}$ everywhere inside the conductor; otherwise, charges would move. In particular, at any point just inside the surface the component of $\vec{E}$ tangent to the surface is zero. It follows that the tangential component of $\vec{E}$ is also zero just *outside* the surface. If it were not, a charge could move around a rectangular path partly inside and partly outside (**Fig. 23.25**) and return to its starting point with a net amount of work having been done on it. This would violate the conservative nature of electrostatic fields, so the tangential component of $\vec{E}$ just outside the surface must be zero at every point on the surface. Thus $\vec{E}$ is perpendicular to the surface at each point, proving our statement.

It also follows that **when all charges are at rest, the entire solid volume of a conductor is at the same potential.** Equation (23.17) states that the potential difference between two points a and b within the conductor's solid volume, $V_a - V_b$, is equal to the line integral $\int_a^b \vec{E} \cdot d\vec{l}$ of the electric field from a to b. Since $\vec{E} = \mathbf{0}$ everywhere inside the conductor, the integral is guaranteed to be zero for any two such points a and b. Hence the potential is the same for any two points within the solid volume of the conductor. We describe this by saying that the solid volume of the conductor is an *equipotential volume.*

We can now prove a theorem that we quoted without proof in Section 22.5. The theorem is as follows: In an electrostatic situation, if a conductor contains a cavity and if no charge is present inside the cavity, then there can be no net charge *anywhere* on the surface of the cavity. This means that if you're inside a charged conducting box, you can safely touch any point on the inside walls of the box without being shocked. To prove this theorem, we first prove that *every point in the cavity is at the same potential.* In **Fig. 23.26** the conducting surface A of the cavity is an equipotential surface, as we have just proved. Suppose point P in the cavity is at a different potential; then we can construct a different equipotential surface B including point P.

Now consider a Gaussian surface, shown in Fig. 23.26, between the two equipotential surfaces. Because of the relationship between $\vec{E}$ and the equipotentials, we know that the field at every point between the equipotentials is from A toward B, or else at every point it is from B toward A, depending on which equipotential surface is at higher potential.

In either case the flux through this Gaussian surface is certainly not zero. But then Gauss's law says that the charge enclosed by the Gaussian surface cannot be zero. This contradicts our initial assumption that there is *no* charge in the cavity. So the potential at P *cannot* be different from that at the cavity wall.

The entire region of the cavity must therefore be at the same potential. But for this to be true, *the electric field inside the cavity must be zero everywhere*. Finally, Gauss's law shows that the electric field at any point on the surface of a conductor is proportional to the surface charge density σ at that point. We conclude that *the surface charge density on the wall of the cavity is zero at every point*. This chain of reasoning may seem tortuous, but it is worth careful study.

CAUTION **Equipotential surfaces vs. Gaussian surfaces** Don't confuse equipotential surfaces with the Gaussian surfaces we encountered in Chapter 22. Gaussian surfaces have relevance only when we are using Gauss's law, and we can choose *any* Gaussian surface that's convenient. We *cannot* choose equipotential surfaces; the shape is determined by the charge distribution.

TEST YOUR UNDERSTANDING OF SECTION 23.4 Would the shapes of the equipotential surfaces in Fig. 23.23 change if the sign of each charge were reversed?

ANSWER

no If the positive charges in Fig. 23.23 were replaced by negative charges, and vice versa, the equipotential surfaces would be the same but the sign of the potential would be reversed. For example, the surfaces in Fig. 23.23b with potential $V = +30\text{ V}$ and $V = -50\text{ V}$ would have potential $V = -30\text{ V}$ and $V = +50\text{ V}$, respectively.

23.5 POTENTIAL GRADIENT

Electric field and potential are closely related. Equation (23.17), restated here, expresses one aspect of that relationship:

$$V_a - V_b = \int_a^b \vec{E} \cdot d\vec{l}$$

If we know $\vec{E}$ at various points, we can use this equation to calculate potential differences. In this section we show how to turn this around; if we know the potential V at various points, we can use it to determine $\vec{E}$. Regarding V as a function of the coordinates (x, y, z) of a point in space, we'll show that the components of $\vec{E}$ are related to the *partial derivatives* of V with respect to x, y, and z.

In Eq. (23.17), $V_a - V_b$ is the potential of a with respect to b—that is, the change of potential encountered on a trip from b to a. We can write this as

$$V_a - V_b = \int_b^a dV = -\int_a^b dV$$

where dV is the infinitesimal change of potential accompanying an infinitesimal element $d\vec{l}$ of the path from b to a. Comparing to Eq. (23.17), we have

$$-\int_a^b dV = \int_a^b \vec{E} \cdot d\vec{l}$$

These two integrals must be equal for *any* pair of limits a and b, and for this to be true the *integrands* must be equal. Thus, for *any* infinitesimal displacement $d\vec{l}$,

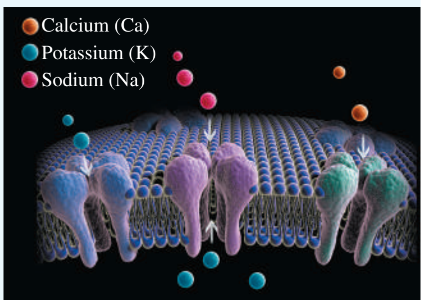
$$-dV = \vec{E} \cdot d\vec{l}$$

To interpret this expression, we write $\vec{E}$ and $d\vec{l}$ in terms of their components: $\vec{E} = \hat{\imath}\, E_x + \hat{\jmath}\, E_y + \hat{k}\, E_z$ and $d\vec{l} = \hat{\imath}\, dx + \hat{\jmath}\, dy + \hat{k}\, dz$. Then

$$-dV = E_x dx + E_y dy + E_z dz$$

Suppose the displacement is parallel to the x-axis, so $dy = dz = 0$. Then $-dV = E_x dx$ or $E_x = -(dV/dx)_{y,\,z\ \text{constant}}$, where the subscript reminds us that only x varies in the

BIO APPLICATION Potential Gradient Across a Cell Membrane The interior of a human cell is at a lower electric potential V than the exterior. (The potential difference when the cell is inactive is about -70 mV in neurons and about -95 mV in skeletal muscle cells.) Hence there is a potential gradient $\vec{\nabla}V$ that points from the *interior* to the *exterior* of the cell membrane, and an electric field $\vec{E} = -\vec{\nabla}V$ that points from the *exterior* to the *interior*. This field affects how ions flow into or out of the cell through special channels in the membrane.



derivative; recall that V is in general a function of x, y, and z. But this is just what is meant by the partial derivative $\partial V/\partial x$. The y- and z-components of $\vec{E}$ are related to the corresponding derivatives of V in the same way, so

Electric field components found from potential:

Each electric field component ...

$$E_x = -\frac{\partial V}{\partial x} \qquad E_y = -\frac{\partial V}{\partial y} \qquad E_z = -\frac{\partial V}{\partial z} \tag{23.19}$$

... equals the negative of the corresponding partial derivative of electric potential function V.

This is consistent with the units of electric field being V/m. In terms of unit vectors we can write

Electric field vector found from potential:

Electric field

$$\vec{E} = -\left(\hat{\imath}\frac{\partial V}{\partial x} + \hat{\jmath}\frac{\partial V}{\partial y} + \hat{k}\frac{\partial V}{\partial z}\right) \tag{23.20}$$

Partial derivatives of electric potential function V

The following operation is called the **gradient** of the function f:

$$\vec{\nabla} f = \left(\hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right) f \tag{23.21}$$

The operator denoted by $\vec{\nabla}$ is called "grad" or "del." Thus in vector notation,

$$\vec{E} = -\vec{\nabla} V \tag{23.22}$$

This is read "$\vec{E}$ is the negative of the gradient of V" or "$\vec{E}$ equals negative grad V." The quantity $\vec{\nabla}V$ is called the *potential gradient*.

At each point, the potential gradient $\vec{\nabla}V$ points in the direction in which V *increases* most rapidly with a change in position. Hence at each point the direction of $\vec{E} = -\vec{\nabla}V$ is the direction in which V *decreases* most rapidly and is always perpendicular to the equipotential surface through the point. This agrees with our observation in Section 23.2 that moving in the direction of the electric field means moving in the direction of decreasing potential.

Equation (23.22) doesn't depend on the particular choice of the zero point for V. If we were to change the zero point, the effect would be to change V at every point by the same amount; the derivatives of V would be the same.

If $\vec{E}$ has a radial component E_r with respect to a point or an axis and r is the distance from the point or axis, the relationship corresponding to Eqs. (23.19) is

$$E_r = -\frac{\partial V}{\partial r} \qquad \text{(radial electric field component)} \tag{23.23}$$

Often we can compute the electric field caused by a charge distribution in either of two ways: directly, by adding the $\vec{E}$ fields of point charges, or by first calculating the potential and then taking its gradient to find the field. The second method is often easier because potential is a *scalar* quantity, requiring at worst the integration of a scalar function. Electric field is a *vector* quantity, requiring computation of components for each element of charge and a separate integration for each component. Thus, quite apart from its fundamental significance, potential offers a very useful computational technique in field calculations. In the next two examples, a knowledge of V is used to find the electric field.

We stress once more that if we know $\vec{E}$ as a function of position, we can calculate V from Eq. (23.17) or (23.18), and if we know V as a function of position, we can calculate $\vec{E}$ from Eq. (23.19), (23.20), or (23.23). Deriving V from $\vec{E}$ requires integration, and deriving $\vec{E}$ from V requires differentiation.

EXAMPLE 23.13 Potential and field of a point charge

From Eq. (23.14) the potential at a radial distance r from a point charge q is $V = q/4\pi\epsilon_0 r$. Find the vector electric field from this expression for V.

IDENTIFY and SET UP This problem uses the general relationship between the electric potential as a function of position and the electric-field vector. By symmetry, the electric field here has only a radial component E_r. We use Eq. (23.23) to find this component.

EXECUTE From Eq. (23.23),

$$E_r = -\frac{\partial V}{\partial r} = -\frac{\partial}{\partial r}\left(\frac{1}{4\pi\epsilon_0}\frac{q}{r}\right) = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}$$

so the vector electric field is

$$\vec{\boldsymbol{E}} = \hat{\boldsymbol{r}}E_r = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}\hat{\boldsymbol{r}}$$

EVALUATE Our result agrees with Eq. (21.7), as it must.

An alternative approach is to ignore the radial symmetry, write the radial distance as $r = \sqrt{x^2 + y^2 + z^2}$, and take the derivatives of V with respect to x, y, and z as in Eq. (23.20). We find

$$\frac{\partial V}{\partial x} = \frac{\partial}{\partial x}\left(\frac{1}{4\pi\epsilon_0}\frac{q}{\sqrt{x^2 + y^2 + z^2}}\right) = -\frac{1}{4\pi\epsilon_0}\frac{qx}{(x^2 + y^2 + z^2)^{3/2}}$$
$$= -\frac{qx}{4\pi\epsilon_0 r^3}$$

and similarly

$$\frac{\partial V}{\partial y} = -\frac{qy}{4\pi\epsilon_0 r^3} \qquad \frac{\partial V}{\partial z} = -\frac{qz}{4\pi\epsilon_0 r^3}$$

Then from Eq. (23.20),

$$\vec{\boldsymbol{E}} = -\left[\hat{\imath}\left(-\frac{qx}{4\pi\epsilon_0 r^3}\right) + \hat{\jmath}\left(-\frac{qy}{4\pi\epsilon_0 r^3}\right) + \hat{\boldsymbol{k}}\left(-\frac{qz}{4\pi\epsilon_0 r^3}\right)\right]$$
$$= \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}\left(\frac{x\hat{\imath} + y\hat{\jmath} + z\hat{\boldsymbol{k}}}{r}\right) = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}\hat{\boldsymbol{r}}$$

This approach gives us the same answer, but with more effort. Clearly it's best to exploit the symmetry of the charge distribution whenever possible.

KEYCONCEPT If you know the electric potential as a function of position in a region of space, you can find the electric field in that region by calculating the negative gradient of the potential.

EXAMPLE 23.14 Potential and field of a ring of charge

In Example 23.11 (Section 23.3) we found that for a ring of charge with radius a and total charge Q, the potential at a point P on the ring's symmetry axis a distance x from the center (see Fig. 23.20) is

$$V = \frac{1}{4\pi\epsilon_0}\frac{Q}{\sqrt{x^2 + a^2}}$$

Find the electric field at P.

IDENTIFY and SET UP We are given V as a function of x along the x-axis, and we wish to find the electric field at a point on this axis. From the symmetry of the charge distribution, the electric field along the symmetry (x-) axis of the ring can have only an x-component. We find it by using the first of Eqs. (23.19).

EXECUTE The x-component of the electric field is

$$E_x = -\frac{\partial V}{\partial x} = \frac{1}{4\pi\epsilon_0}\frac{Qx}{(x^2 + a^2)^{3/2}}$$

EVALUATE This agrees with our result in Example 21.9.

CAUTION Don't use expressions where they don't apply In this example, V is not a function of y or z on the ring axis, so $\partial V/\partial y = \partial V/\partial z = 0$ and $E_y = E_z = 0$. But that does not mean that it's true *everywhere;* our expressions for V and E_x are valid *on the ring axis only.* If we had an expression for V valid at *all* points in space, we could use it to find the components of $\vec{\boldsymbol{E}}$ at any point by using Eqs. (23.19).

KEYCONCEPT If you know the electric potential as a function of position only along a line, you can find the component of the electric field parallel to that line by calculating the negative partial derivative of the potential with respect to position on that line. You cannot find the electric field at points off the line in this way.

TEST YOUR UNDERSTANDING OF SECTION 23.5 In a certain region of space the potential is given by $V = A + Bx + Cy^3 + Dxy$, where A, B, C, and D are positive constants. Which of these statements about the electric field $\vec{\boldsymbol{E}}$ in this region of space is correct? (There may be more than one correct answer.) (i) Increasing the value of A will increase the value of $\vec{\boldsymbol{E}}$ at all points; (ii) increasing the value of A will decrease the value of $\vec{\boldsymbol{E}}$ at all points; (iii) $\vec{\boldsymbol{E}}$ has no z-component; (iv) the electric field is zero at the origin ($x = 0$, $y = 0$, $z = 0$).

ANSWER

(iii) From Eqs. (23.19), the components of the electric field are $E_x = -\partial V/\partial x = -(B + Dy)$, $E_y = -\partial V/\partial y = -(3Cy^2 + Dx)$, and $E_z = -\partial V/\partial z = 0$. The value of A has no effect, which means that we can add a constant to the electric potential at all points without changing $\vec{\boldsymbol{E}}$ or the potential difference between two points. The potential does not depend on z, so the z-component of $\vec{\boldsymbol{E}}$ is zero. Note that at the origin the electric field is not zero because it has a nonzero x-component: $E_x = -B$, $E_y = 0$, $E_z = 0$.

CHAPTER 23 SUMMARY

Electric potential energy: The electric force caused by any collection of charges at rest is a conservative force. The work W done by the electric force on a charged particle moving in an electric field can be represented by the change in a potential-energy function U.

The electric potential energy for two point charges q and q_0 depends on their separation r. The electric potential energy for a charge q_0 in the presence of a collection of charges q_1, q_2, q_3 depends on the distance from q_0 to each of these other charges. (See Examples 23.1 and 23.2.)

$$W_{a\to b} = U_a - U_b \tag{23.2}$$

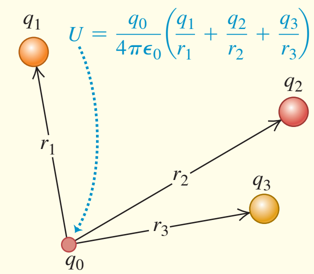
$$U = \frac{1}{4\pi\epsilon_0}\frac{qq_0}{r} \tag{23.9}$$

(two point charges)

$$U = \frac{q_0}{4\pi\epsilon_0}\left(\frac{q_1}{r_1} + \frac{q_2}{r_2} + \frac{q_3}{r_3} + \cdots\right)$$

$$= \frac{q_0}{4\pi\epsilon_0}\sum_i \frac{q_i}{r_i} \tag{23.10}$$

(q_0 in presence of other point charges)



Electric potential: Potential, denoted by V, is potential energy per unit charge. The potential difference between two points equals the amount of work per charge that would be required to move a positive test charge between those points. The potential V due to a quantity of charge can be calculated by summing (if the charge is a collection of point charges) or by integrating (if the charge is a distribution). (See Examples 23.3, 23.4, 23.5, 23.7, 23.11, and 23.12.)

The potential difference between two points a and b, also called the potential of a with respect to b, is given by the line integral of $\vec{E}$. The potential at a given point can be found by first finding $\vec{E}$ and then carrying out this integral. (See Examples 23.6, 23.8, 23.9, and 23.10.)

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{r} \tag{23.14}$$

(due to a point charge)

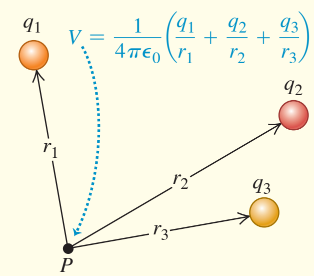
$$V = \frac{1}{4\pi\epsilon_0}\sum_i \frac{q_i}{r_i} \tag{23.15}$$

(due to a collection of point charges)

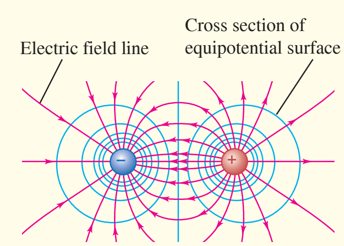
$$V = \frac{1}{4\pi\epsilon_0}\int \frac{dq}{r} \tag{23.16}$$

(due to a charge distribution)

$$V_a - V_b = \int_a^b \vec{E}\cdot d\vec{l} = \int_a^b E\cos\phi\, dl \tag{23.17}$$



Equipotential surfaces: An equipotential surface is a surface on which the potential has the same value at every point. At a point where a field line crosses an equipotential surface, the two are perpendicular. When all charges are at rest, the surface of a conductor is always an equipotential surface and all points in the interior of a conductor are at the same potential. When a cavity within a conductor contains no charge, the entire cavity is an equipotential region and there is no surface charge anywhere on the surface of the cavity.



Finding electric field from electric potential: If the potential V is known as a function of the coordinates x, y, and z, the components of electric field $\vec{E}$ at any point are given by partial derivatives of V. (See Examples 23.13 and 23.14.)

$$E_x = -\frac{\partial V}{\partial x} \quad E_y = -\frac{\partial V}{\partial y} \quad E_z = -\frac{\partial V}{\partial z} \tag{23.19}$$

$$\vec{E} = -\left(\hat{\imath}\frac{\partial V}{\partial x} + \hat{\jmath}\frac{\partial V}{\partial y} + \hat{k}\frac{\partial V}{\partial z}\right) \tag{23.20}$$

(vector form)

GUIDED PRACTICE

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 23.1 and 23.2 (Section 23.1) before attempting these problems.

VP23.2.1 An electron (charge $-e = -1.60 \times 10^{-19}$ C) is at rest at a distance 1.00×10^{-10} m from the center of a nucleus of lead (charge $+82e = +1.31 \times 10^{-17}$ C). You may treat both objects as point charges. (a) How much work would you have to do to move the electron to a distance 5.00×10^{-10} m from the center of the lead nucleus? (b) If you release the electron from the position in (a), what will be its kinetic energy when it is a distance 8.00×10^{-12} m from the center of the lead nucleus? (You can assume that the lead nucleus remains at rest, since it is much more massive than the electron.)

VP23.2.2 In a nuclear physics experiment, a proton (mass 1.67×10^{-27} kg, charge $+e = +1.60 \times 10^{-19}$ C) is fired directly at a target nucleus of unknown charge. (You can treat both objects as point charges, and assume that the nucleus remains at rest.) When it is far from its target, the proton has speed 2.50×10^{6} m/s. The proton comes momentarily to rest at a distance 5.29×10^{-13} m from the center of the target nucleus, then flies back in the direction from which it came. (a) What is the electric potential energy of the proton and nucleus when they are 5.29×10^{-13} m apart? (b) What is the charge of the target nucleus?

VP23.2.3 Two point charges are at fixed positions on the y-axis: $q_1 = +e$ at $y = 0$ and $q_2 = -e$ at $y = a$. Find (a) the work you must do to bring a third charge $q_3 = -e$ from infinity to being at rest at $y = -2a$ and (b) the total potential energy of the system of three charges.

VP23.2.4 A point charge $q_1 = +5.00$ nC is at the fixed position $x = 0$, $y = 0$, $z = 0$. You find that you must do 8.10×10^{-6} J of work to bring a second point charge from infinity to the position $x = +4.00$ cm, $y = 0$, $z = 0$. (a) What is the value of the second charge? (b) Once the second charge is in place, how much additional work would you have to do to bring a third point charge $q_3 = +2.00$ nC from infinity to the position $x = 0$, $y = +3.00$ cm, $z = 0$?

Be sure to review EXAMPLES 23.3, 23.4, 23.5, and 23.7 (Section 23.2) before attempting these problems.

VP23.7.1 In a certain region of space the electric field is uniform and given by $\vec{E} = (5.00 \times 10^{2}\ \text{V/m})\hat{\imath}$. If the electric potential at the point $x = 0$, $y = 0$, $z = 0$ is equal to V_0, find the potential difference $V_0 - V_P$ for each of the following points P: (a) $x = +5.00$ cm, $y = 0$, $z = 0$; (b) $x = +3.00$ cm, $y = +4.00$ cm, $z = 0$; (c) $x = 0$, $y = +5.00$ cm, $z = 0$; (d) $x = -5.00$ cm, $y = 0$, $z = 0$.

VP23.7.2 You have two identical point charges of +6.0 nC each, one at the position $x = 3.0$ cm, $y = 0$, $z = 0$ and the other at $x = -3.0$ cm, $y = 0$, $z = 0$. Find the electric potential due to these charges at (a) $x = 0$, $y = 0$, $z = 0$; (b) $x = 0$, $y = 4.0$ cm, $z = 0$; (c) $x = 8.0$ cm, $y = 0$, $z = 0$. Take $V = 0$ at infinity. (d) At which (if any) of these points is the electric *field* equal to zero? If the field is zero at a certain point, is the potential necessarily zero there as well?

VP23.7.3 One point charge of +5.90 nC is at the position $x = 3.00$ cm, $y = 0$, $z = 0$, and one point charge of −5.90 nC is at $x = -3.00$ cm, $y = 0$, $z = 0$. Find the electric potential due to these charges at (a) $x = 3.00$ cm, $y = 8.00$ cm, $z = 0$; (b) $x = -3.00$ cm, $y = 8.00$ cm, $z = 0$. Take $V = 0$ at infinity. (c) A proton (mass 1.67×10^{-27} kg, charge $+e = +1.60 \times 10^{-19}$ C) follows a path from point (b) to point (a). If its speed at (b) is 3.00×10^{5} m/s, what is its speed at (a)?

VP23.7.4 In Example 22.9 (Section 22.4) we found that the electric field *inside* a uniformly charged sphere with positive charge Q and radius R points radially outward from the center of the sphere and has magnitude $E = Qr/4\pi\epsilon_0 R^3$ at a distance r from the center. (a) By integrating the electric field, find the potential difference between the center of the sphere and its surface. (b) Which is at higher potential: the center or the surface?

Be sure to review EXAMPLES 23.8, 23.9, 23.10, 23.11, and 23.12 (Section 23.3) before attempting these problems.

VP23.12.1 A conducting sphere of radius R carries positive charge q. Calculate the amount of work that would be required to move a small positive test charge q_0 slowly (a) from $r = 5R$ to $r = 3R$; (b) from $r = 2R$ to $r = 7R$; (c) from $r = 4R$ to $r = R/2$. In each case assume that the presence of q_0 has no effect on how the charge q is distributed over the sphere.

VP23.12.2 For the two oppositely charged parallel plates in Fig. 23.18, the potential difference between the plates is $V_{ab} = 24.0$ V, $V = 0$ at the bottom plate, and the distance between the plates is $d =$ 4.50 mm. (a) What is the electric potential at a point 3.00 mm above the lower plate? (b) What is the potential energy for a dust particle of mass 5.00×10^{-9} kg and charge +2.00 nC at the position in part (a)? (c) If the particle in part (b) is released from rest, will it move toward the upper plate or the lower plate? What will be its speed when it hits that plate?

VP23.12.3 The uniformly charged ring shown in Fig. 23.20 has charge +5.00 nC and radius 2.50 cm. A point charge $q = +3.00$ nC of mass 4.00×10^{-9} kg is on the ring axis 8.00 cm from the center of the ring and is moving toward the center at 60.0 m/s. Take $V = 0$ at infinity. What is the potential energy of the point charge (a) at 8.00 cm from the ring's center and (b) at the ring's center? (c) What are the kinetic energy and speed of the point charge when it reaches the ring's center?

VP23.12.4 Positive charge Q is distributed uniformly along a rod of length L that lies along the x-axis from $x = L$ to $x = 2L$. (a) How much charge is contained within a segment of the rod of length dx? (b) Integrate to find the total electric potential at the origin ($x = 0$) due to the rod.

BRIDGING PROBLEM A Point Charge and a Line of Charge

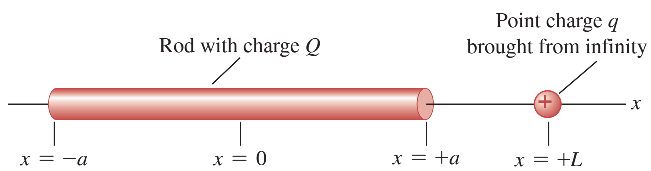
Positive electric charge Q is distributed uniformly along a thin rod of length $2a$. The rod lies along the x-axis between $x = -a$ and $x = +a$ (**Fig. 23.27**). Calculate how much work you must do to bring a positive point charge q from infinity to the point $x = +L$ on the x-axis, where $L > a$.

SOLUTION GUIDE

IDENTIFY and SET UP

1. In this problem you must first calculate the potential V at $x = +L$ due to the charged rod. You can then find the change in potential energy involved in bringing the point charge q from infinity (where $V = 0$) to $x = +L$.
2. To find V, divide the rod into infinitesimal segments of length dx'. How much charge is on such a segment? Consider one such segment located at $x = x'$, where $-a \le x' \le a$. What is the potential dV at $x = +L$ due to this segment?
3. The total potential at $x = +L$ is the integral of dV, including contributions from all of the segments for x' from $-a$ to $+a$. Set up this integral.

Figure 23.27 How much work must you do to bring point charge q in from infinity?



EXECUTE

4. Integrate your expression from step 3 to find the potential V at $x = +L$. A simple, standard substitution will do the trick; use a table of integrals only as a last resort.
5. Use your result from step 4 to find the potential energy for a point charge q placed at $x = +L$.
6. Use your result from step 5 to find the work you must do to bring the point charge from infinity to $x = +L$.

EVALUATE

7. What does your result from step 5 become in the limit $a \to 0$? Does this make sense?
8. Suppose the point charge q were negative rather than positive. How would this affect your result in step 4? In step 5?

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q23.1 A student asked, "Since electrical potential is always proportional to potential energy, why bother with the concept of potential at all?" How would you respond?

Q23.2 The potential (relative to a point at infinity) midway between two charges of equal magnitude and opposite sign is zero. Is it possible to bring a test charge from infinity to this midpoint in such a way that no work is done in any part of the displacement? If so, describe how it can be done. If it is not possible, explain why.

Q23.3 Is it possible to have an arrangement of two point charges separated by a finite distance such that the electric potential energy of the arrangement is the same as if the two charges were infinitely far apart? Why or why not? What if there are three charges? Explain.

Q23.4 Since potential can have any value you want depending on the choice of the reference level of zero potential, how does a voltmeter know what to read when you connect it between two points?

Q23.5 If $\vec{E}$ is zero everywhere along a certain path that leads from point A to point B, what is the potential difference between those two points? Does this mean that $\vec{E}$ is zero everywhere along *any* path from A to B? Explain.

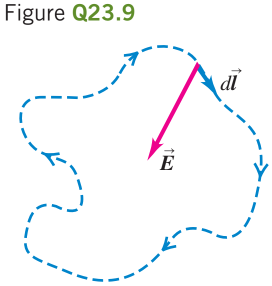
Q23.6 If $\vec{E}$ is zero throughout a certain region of space, is the potential necessarily also zero in this region? Why or why not? If not, what *can* be said about the potential?

Q23.7 Which way do electric field lines point, from high to low potential or from low to high? Explain.

Q23.8 (a) Take $V = 0$ at infinity. If the potential is zero at a point, is the electric field necessarily zero at that point? (b) If the electric field is zero at a point, is the potential necessarily zero there? Prove your answers, using simple examples.

Q23.9 If you carry out the integral of the electric field $\int \vec{E} \cdot d\vec{l}$ for a *closed* path like that shown in **Fig. Q23.9**, the integral will *always* be equal to zero, independent of the shape of the path and independent of where charges may be located relative to the path. Explain why.

Figure Q23.9



Q23.10 The potential difference between the two terminals of an AA battery (used in flashlights and portable stereos) is 1.5 V. If two AA batteries are placed end to end with the positive terminal of one battery touching the negative terminal of the other, what is the potential difference between the terminals at the exposed ends of the combination? What if the two positive terminals are touching each other? Explain your reasoning.

Q23.11 It is easy to produce a potential difference of several thousand volts between your body and the floor by scuffing your shoes across a nylon carpet. When you touch a metal doorknob, you get a mild shock. Yet contact with a power line of comparable voltage would probably be fatal. Why is there a difference?

Q23.12 If the electric potential at a single point is known, can $\vec{E}$ at that point be determined? If so, how? If not, why not?

Q23.13 Because electric field lines and equipotential surfaces are always perpendicular, two equipotential surfaces can never cross; if they did, the direction of $\vec{E}$ would be ambiguous at the crossing points. Yet two equipotential surfaces appear to cross at the center of Fig. 23.23c. Explain why there is no ambiguity about the direction of $\vec{E}$ in this particular case.

Q23.14 A uniform electric field is directed due east. Point B is 2.00 m west of point A, point C is 2.00 m east of point A, and point D is 2.00 m south of A. For each point, B, C, and D, is the potential at that point larger, smaller, or the same as at point A? Give the reasoning behind your answers.

Q23.15 We often say that if point A is at a higher potential than point B, A is at positive potential and B is at negative potential. Does it necessarily follow that a point at positive potential is positively charged, or that a point at negative potential is negatively charged? Illustrate your answers with clear, simple examples.

Q23.16 A conducting sphere is to be charged by bringing in positive charge a little at a time until the total charge is Q. The total work required for this process is alleged to be proportional to Q^2. Is this correct? Why or why not?

Q23.17 In electronics it is customary to define the potential of ground (thinking of the earth as a large conductor) as zero. Is this consistent with the fact that the earth has a net electric charge that is not zero?

Q23.18 A conducting sphere is placed between two charged parallel plates such as those shown in Fig. 23.2. Does the electric field inside the sphere depend on precisely where between the plates the sphere is placed? What about the electric potential inside the sphere? Do the answers to these questions depend on whether or not there is a net charge on the sphere? Explain your reasoning.

Q23.19 A conductor that carries a net charge Q has a hollow, empty cavity in its interior. Does the potential vary from point to point within the material of the conductor? What about within the cavity? How does the potential inside the cavity compare to the potential within the material of the conductor?

Q23.20 A high-voltage dc power line falls on a car, so the entire metal body of the car is at a potential of 10,000 V with respect to the ground. What happens to the occupants (a) when they are sitting in the car and (b) when they step out of the car? Explain your reasoning.

Q23.21 When a thunderstorm is approaching, sailors at sea sometimes observe a phenomenon called "St. Elmo's fire," a bluish flickering light at the tips of masts. What causes this? Why does it occur at the tips of masts? Why is the effect most pronounced when the masts are wet? (*Hint:* Seawater is a good conductor of electricity.)

Q23.22 A positive point charge is placed near a very large conducting plane. A professor of physics asserted that the field caused by this configuration is the same as would be obtained by removing the plane and placing a negative point charge of equal magnitude in the mirror-image position behind the initial position of the plane. Is this correct? Why or why not? (*Hint:* Inspect Fig. 23.23b.)

EXERCISES

Section 23.1 Electric Potential Energy

23.1 •• A point charge $q_1 = +2.40\ \mu\text{C}$ is held stationary at the origin. A second point charge $q_2 = -4.30\ \mu\text{C}$ moves from the point $x = 0.150$ m, $y = 0$ to the point $x = 0.250$ m, $y = 0.250$ m. How much work is done by the electric force on q_2?

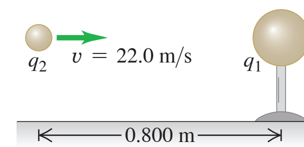
23.2 • A point charge q_1 is held stationary at the origin. A second charge q_2 is placed at point a, and the electric potential energy of the pair of charges is $+5.4 \times 10^{-8}$ J. When the second charge is moved to point b, the electric force on the charge does -1.9×10^{-8} J of work. What is the electric potential energy of the pair of charges when the second charge is at point b?

23.3 •• **Energy of the Nucleus.** How much work is needed to assemble an atomic nucleus containing three protons (such as Li) if we model it as an equilateral triangle of side 2.00×10^{-15} m with a proton at each vertex? Assume the protons started from very far away.

23.4 •• (a) How much work would it take to push two protons very slowly from a separation of 2.00×10^{-10} m (a typical atomic distance) to 3.00×10^{-15} m (a typical nuclear distance)? (b) If the protons are both released from rest at the closer distance in part (a), how fast are they moving when they reach their original separation?

23.5 •• A small metal sphere, carrying a net charge of $q_1 = -2.80\ \mu\text{C}$, is held in a stationary position by insulating supports. A second small metal sphere, with a net charge of $q_2 = -7.80\ \mu\text{C}$ and mass 1.50 g, is projected toward q_1. When the two spheres are 0.800 m apart, q_2 is moving toward q_1 with speed 22.0 m/s (**Fig. E23.5**). Assume that the two spheres can be treated as point charges. You can ignore the force of gravity. (a) What is the speed of q_2 when the spheres are 0.400 m apart? (b) How close does q_2 get to q_1?

Figure **E23.5**



23.6 •• **BIO Energy of DNA Base Pairing.** (See Exercise 21.18.) (a) Calculate the electric potential energy of the adenine–thymine bond, using the same combinations of molecules (O–H–N and N–H–N) as in Exercise 21.18. (b) Compare this energy with the potential energy of the proton–electron pair in the hydrogen atom.

23.7 •• Two protons, starting several meters apart, are aimed directly at each other with speeds of 2.00×10^5 m/s, measured relative to the earth. Find the maximum electric force that these protons will exert on each other.

23.8 •• Three equal $1.20\ \mu\text{C}$ point charges are placed at the corners of an equilateral triangle with sides 0.400 m long. What is the potential energy of the system? (Take as zero the potential energy of the three charges when they are infinitely far apart.)

23.9 •• Two protons are released from rest when they are 0.750 nm apart. (a) What is the maximum speed they will reach? When does this speed occur? (b) What is the maximum acceleration they will achieve? When does this acceleration occur?

23.10 •• Four electrons are located at the corners of a square 10.0 nm on a side, with an alpha particle at its midpoint. How much work is needed to move the alpha particle to the midpoint of one of the sides of the square?

Section 23.2 Electric Potential

23.11 •• **CALC** Points a and b lie in a region where the y-component of the electric field is $E_y = \alpha + \beta/y^2$. The constants in this expression have the values $\alpha = 600$ N/C and $\beta = 5.00\ \text{N} \cdot \text{m}^2/\text{C}$. Points a and b are on the $+y$-axis. Point a is at $y = 2.00$ cm and point b is at $y = 3.00$ cm. What is the potential difference $V_a - V_b$ between these two points and which point, a or b, is at higher potential?

23.12 • An object with charge $q = -6.00 \times 10^{-9}$ C is placed in a region of uniform electric field and is released from rest at point A. After the charge has moved to point B, 0.500 m to the right, it has kinetic energy 3.00×10^{-7} J. (a) If the electric potential at point A is +30.0 V, what is the electric potential at point B? (b) What are the magnitude and direction of the electric field?

23.13 • A small particle has charge $-5.00\ \mu\text{C}$ and mass 2.00×10^{-4} kg. It moves from point A, where the electric potential is $V_A = +200$ V, to point B, where the electric potential is $V_B = +800$ V. The electric force is the only force acting on the particle. The particle has speed 5.00 m/s at point A. What is its speed at point B? Is it moving faster or slower at B than at A? Explain.

23.14 • A particle with charge +4.20 nC is in a uniform electric field $\vec{E}$ directed to the left. The charge is released from rest and moves to the left; after it has moved 6.00 cm, its kinetic energy is $+2.20 \times 10^{-6}$ J. What are (a) the work done by the electric force, (b) the potential of the starting point with respect to the end point, and (c) the magnitude of $\vec{E}$?

23.15 • A charge of 28.0 nC is placed in a uniform electric field that is directed vertically upward and has a magnitude of 4.00×10^4 V/m. What work is done by the electric force when the charge moves (a) 0.450 m to the right; (b) 0.670 m upward; (c) 2.60 m at an angle of 45.0° downward from the horizontal?

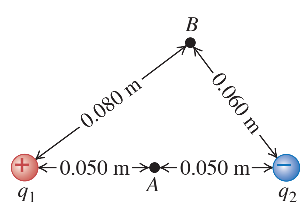
23.16 • Two stationary point charges +3.00 nC and +2.00 nC are separated by a distance of 50.0 cm. An electron is released from rest at a point midway between the two charges and moves along the line connecting the two charges. What is the speed of the electron when it is 10.0 cm from the +3.00 nC charge?

23.17 •• Point charges $q_1 = +2.00\ \mu C$ and $q_2 = -2.00\ \mu C$ are placed at adjacent corners of a square for which the length of each side is 3.00 cm. Point a is at the center of the square, and point b is at the empty corner closest to q_2. Take the electric potential to be zero at a distance far from both charges. (a) What is the electric potential at point a due to q_1 and q_2? (b) What is the electric potential at point b? (c) A point charge $q_3 = -5.00\ \mu C$ moves from point a to point b. How much work is done on q_3 by the electric forces exerted by q_1 and q_2? Is this work positive or negative?

23.18 • Two point charges of equal magnitude Q are held a distance d apart. Consider only points on the line passing through both charges; take $V = 0$ at infinity. (a) If the two charges have the same sign, find the location of all points (if there are any) at which (i) the potential is zero (is the electric field zero at these points?), and (ii) the electric field is zero (is the potential zero at these points?). (b) Repeat part (a) for two point charges having opposite signs.

23.19 • Two point charges $q_1 = +2.40$ nC and $q_2 = -6.50$ nC are 0.100 m apart. Point A is midway between them; point B is 0.080 m from q_1 and 0.060 m from q_2 (**Fig. E23.19**). Take the electric potential to be zero at infinity. Find (a) the potential at point A; (b) the potential at point B; (c) the work done by the electric field on a charge of 2.50 nC that travels from point B to point A.

Figure **E23.19**



23.20 •• (a) An electron is to be accelerated from 3.00×10^6 m/s to 8.00×10^6 m/s. Through what potential difference must the electron pass to accomplish this? (b) Through what potential difference must the electron pass if it is to be slowed from 8.00×10^6 m/s to a halt?

23.21 •• Points A and B lie within a region of space where there is a uniform electric field that has no x- or z-component; only the y-component E_y is nonzero. Point A is at $y = 8.00$ cm and point B is at $y = 15.0$ cm. The potential difference between B and A is $V_B - V_A = +12.0$ V, so point B is at higher potential than point A. (a) Is E_y positive or negative? (b) What is the magnitude of the electric field? (c) Point C has coordinates $x = 5.00$ cm, $y = 5.00$ cm. What is the potential difference between points B and C?

23.22 • At a certain distance from a point charge, the potential and electric-field magnitude due to that charge are 4.98 V and 16.2 V/m, respectively. (Take $V = 0$ at infinity.) (a) What is the distance to the point charge? (b) What is the magnitude of the charge? (c) Is the electric field directed toward or away from the point charge?

23.23 • A uniform electric field has magnitude E and is directed in the negative x-direction. The potential difference between point a (at $x = 0.60$ m) and point b (at $x = 0.90$ m) is 240 V. (a) Which point, a or b, is at the higher potential? (b) Calculate the value of E. (c) A negative point charge $q = -0.200\ \mu C$ is moved from b to a. Calculate the work done on the point charge by the electric field.

23.24 •• A small sphere with charge $q = -5.00\ \mu C$ is moving in a uniform electric field that has no y- or z-component. The only force on the sphere is the force exerted by the electric field. Point A is on the x-axis at $x = -0.400$ m, and point B is at the origin. At point A the sphere has kinetic energy $K_A = 8.00 \times 10^{-4}$ J, and at point B its kinetic energy is $K_B = 3.00 \times 10^{-4}$ J. (a) What is the potential difference $V_{AB} = V_A - V_B$? Which point, A or B, is at higher potential? (b) What are the magnitude and direction of the electric field?

23.25 •• Identical point charges q_1 and q_2 each have positive charge $+6.00\ \mu C$. Charge q_1 is held fixed on the x-axis at $x = +0.400$ m, and q_2 is held fixed on the x-axis at $x = -0.400$ m. A small sphere has charge $Q = -0.200\ \mu C$ and mass 12.0 g. The sphere is initially very far from the origin. It is released from rest and moves along the y-axis toward the origin. (a) As the sphere moves from very large y to $y = 0$, how much work is done on it by the resultant force exerted by q_1 and q_2? (b) If the only force acting on the sphere is the force exerted by the point charges, what is its speed when it reaches the origin?

Section 23.3 Calculating Electric Potential

23.26 •• A solid conducting sphere of radius 5.00 cm carries a net charge. To find the value of the charge, you measure the potential difference $V_{AB} = V_A - V_B$ between point A, which is 8.00 cm from the center of the sphere, and point B, which is a distance r from the center of the sphere. You repeat these measurements for several values of $r > 8.00$ cm. When you plot your data as V_{AB} versus $1/r$, the values lie close to a straight line with slope -18.0 V·m. What does your data give for the net charge on the sphere? Is the net charge positive or negative?

23.27 •• A thin spherical shell with radius $R_1 = 3.00$ cm is concentric with a larger thin spherical shell with radius $R_2 = 5.00$ cm. Both shells are made of insulating material. The smaller shell has charge $q_1 = +6.00$ nC distributed uniformly over its surface, and the larger shell has charge $q_2 = -9.00$ nC distributed uniformly over its surface. Take the electric potential to be zero at an infinite distance from both shells. (a) What is the electric potential due to the two shells at the following distance from their common center: (i) $r = 0$; (ii) $r = 4.00$ cm; (iii) $r = 6.00$ cm? (b) What is the magnitude of the potential difference between the surfaces of the two shells? Which shell is at higher potential: the inner shell or the outer shell?

23.28 • A total electric charge of 3.50 nC is distributed uniformly over the surface of a metal sphere with a radius of 24.0 cm. If the potential is zero at a point at infinity, find the value of the potential at the following distances from the center of the sphere: (a) 48.0 cm; (b) 24.0 cm; (c) 12.0 cm.

23.29 •• A uniformly charged, thin ring has radius 15.0 cm and total charge +24.0 nC. An electron is placed on the ring's axis a distance 30.0 cm from the center of the ring and is constrained to stay on the axis of the ring. The electron is then released from rest. (a) Describe the subsequent motion of the electron. (b) Find the speed of the electron when it reaches the center of the ring.

23.30 • A solid conducting sphere has net positive charge and radius $R = 0.400$ m. At a point 1.20 m from the center of the sphere, the electric potential due to the charge on the sphere is 24.0 V. Assume that $V = 0$ at an infinite distance from the sphere. What is the electric potential at the center of the sphere?

23.31 • Charge $Q = 5.00\ \mu C$ is distributed uniformly over the volume of an insulating sphere that has radius $R = 12.0$ cm. A small sphere with charge $q = +3.00\ \mu C$ and mass 6.00×10^{-5} kg is projected toward the center of the large sphere from an initial large distance. The large sphere is held at a fixed position and the small sphere can be treated as a point charge. What minimum speed must the small sphere have in order to come within 8.00 cm of the surface of the large sphere?

23.32 •• An infinitely long line of charge has linear charge density 5.00×10^{-12} C/m. A proton (mass 1.67×10^{-27} kg, charge $+1.60 \times 10^{-19}$ C) is 18.0 cm from the line and moving directly toward the line at 3.50×10^3 m/s. (a) Calculate the proton's initial kinetic energy. (b) How close does the proton get to the line of charge?

23.33 •• A very long insulating cylindrical shell of radius 6.00 cm carries charge of linear density 8.50 μC/m spread uniformly over its outer surface. What would a voltmeter read if it were connected between (a) the surface of the cylinder and a point 4.00 cm above the surface, and (b) the surface and a point 1.00 cm from the central axis of the cylinder?

23.34 •• A very long insulating cylinder of charge of radius 2.50 cm carries a uniform linear density of 15.0 nC/m. If you put one probe of a voltmeter at the surface, how far from the surface must the other probe be placed so that the voltmeter reads 175 V?

23.35 •• A very small sphere with positive charge $q = +8.00\ \mu$C is released from rest at a point 1.50 cm from a very long line of uniform linear charge density $\lambda = +3.00\ \mu$C/m. What is the kinetic energy of the sphere when it is 4.50 cm from the line of charge if the only force on it is the force exerted by the line of charge?

23.36 • **CP** Two large, parallel conducting plates carrying opposite charges of equal magnitude are separated by 2.20 cm. (a) If the surface charge density for each plate has magnitude 47.0 nC/m^2, what is the magnitude of $\vec{E}$ in the region between the plates? (b) What is the potential difference between the two plates? (c) If the separation between the plates is doubled while the surface charge density is kept constant at the value in part (a), what happens to the magnitude of the electric field and to the potential difference?

23.37 • Two large, parallel, metal plates carry opposite charges of equal magnitude. They are separated by 45.0 mm, and the potential difference between them is 360 V. (a) What is the magnitude of the electric field (assumed to be uniform) in the region between the plates? (b) What is the magnitude of the force this field exerts on a particle with charge +2.40 nC? (c) Use the results of part (b) to compute the work done by the field on the particle as it moves from the higher-potential plate to the lower. (d) Compare the result of part (c) to the change of potential energy of the same charge, computed from the electric potential.

23.38 • **BIO** **Electrical Sensitivity of Sharks.** Certain sharks can detect an electric field as weak as 1.0 μV/m. To grasp how weak this field is, if you wanted to produce it between two parallel metal plates by connecting an ordinary 1.5 V AA battery across these plates, how far apart would the plates have to be?

23.39 • The electric field at the surface of a charged, solid, copper sphere with radius 0.200 m is 3800 N/C, directed toward the center of the sphere. What is the potential at the center of the sphere, if we take the potential to be zero infinitely far from the sphere?

23.40 •• (a) How much excess charge must be placed on a copper sphere 25.0 cm in diameter so that the potential of its center is 3.75 kV? Take the point where $V = 0$ to be infinitely far from the sphere. (b) What is the potential of the sphere's surface?

Section 23.4 Equipotential Surfaces and Section 23.5 Potential Gradient

23.41 •• **CALC** A metal sphere with radius r_a is supported on an insulating stand at the center of a hollow, metal, spherical shell with radius r_b. There is charge $+q$ on the inner sphere and charge $-q$ on the outer spherical shell. (a) Calculate the potential $V(r)$ for (i) $r < r_a$; (ii) $r_a < r < r_b$; (iii) $r > r_b$. (*Hint:* The net potential is the sum of the potentials due to the individual spheres.) Take V to be zero when r is infinite. (b) Show that the potential of the inner sphere with respect to the outer is

$$V_{ab} = \frac{q}{4\pi\epsilon_0}\left(\frac{1}{r_a} - \frac{1}{r_b}\right)$$

(c) Use Eq. (23.23) and the result from part (a) to show that the electric field at any point between the spheres has magnitude

$$E(r) = \frac{V_{ab}}{(1/r_a - 1/r_b)}\frac{1}{r^2}$$

(d) Use Eq. (23.23) and the result from part (a) to find the electric field at a point outside the larger sphere at a distance r from the center, where $r > r_b$. (e) Suppose the charge on the outer sphere is not $-q$ but a negative charge of different magnitude, say $-Q$. Show that the answers for parts (b) and (c) are the same as before but the answer for part (d) is different.

23.42 • A very large plastic sheet carries a uniform charge density of -6.00 nC/m^2 on one face. (a) As you move away from the sheet along a line perpendicular to it, does the potential increase or decrease? How do you know, without doing any calculations? Does your answer depend on where you choose the reference point for potential? (b) Find the spacing between equipotential surfaces that differ from each other by 1.00 V. What type of surfaces are these?

23.43 • **CALC** In a certain region of space, the electric potential is $V(x, y, z) = Axy - Bx^2 + Cy$, where A, B, and C are positive constants. (a) Calculate the x-, y-, and z-components of the electric field. (b) At which points is the electric field equal to zero?

23.44 • **CALC** In a certain region of space the electric potential is given by $V = +Ax^2y - Bxy^2$, where $A = 5.00$ V/m^3 and $B = 8.00$ V/m^3. Calculate the magnitude and direction of the electric field at the point in the region that has coordinates $x = 2.00$ m, $y = 0.400$ m, and $z = 0$.

23.45 • A metal sphere with radius $r_a = 1.20$ cm is supported on an insulating stand at the center of a hollow, metal, spherical shell with radius $r_b = 9.60$ cm. Charge $+q$ is put on the inner sphere and charge $-q$ on the outer spherical shell. The magnitude of q is chosen to make the potential difference between the spheres 500 V, with the inner sphere at higher potential. (a) Use the result of Exercise 23.41(b) to calculate q. (b) With the help of the result of Exercise 23.41(a), sketch the equipotential surfaces that correspond to 500, 400, 300, 200, 100, and 0 V. (c) In your sketch, show the electric field lines. Are the electric field lines and equipotential surfaces mutually perpendicular? Are the equipotential surfaces closer together when the magnitude of $\vec{E}$ is largest?

PROBLEMS

23.46 • **CP** A point charge $q_1 = +5.00\ \mu$C is held fixed in space. From a horizontal distance of 6.00 cm, a small sphere with mass 4.00×10^{-3} kg and charge $q_2 = +2.00\ \mu$C is fired toward the fixed charge with an initial speed of 40.0 m/s. Gravity can be neglected. What is the acceleration of the sphere at the instant when its speed is 25.0 m/s?

23.47 ••• A point charge $q_1 = 4.00$ nC is placed at the origin, and a second point charge $q_2 = -3.00$ nC is placed on the x-axis at $x = +20.0$ cm. A third point charge $q_3 = 2.00$ nC is to be placed on the x-axis between q_1 and q_2. (Take as zero the potential energy of the three charges when they are infinitely far apart.) (a) What is the potential energy of the system of the three charges if q_3 is placed at $x = +10.0$ cm? (b) Where should q_3 be placed to make the potential energy of the system equal to zero?

23.48 •• A point charge +8.00 nC is on the $-x$-axis at $x = -0.200$ m, and a point charge -4.00 nC is on the $+x$-axis at $x = 0.200$ m. (a) In addition to $x = \pm\infty$, at what point on the x-axis is the resultant field of the two charges equal to zero? (b) Let $V = 0$ at $x = \pm\infty$. At what two other points on the x-axis is the total electric potential due to the two charges equal to zero? (c) Is $E = 0$ at either of the points in part (b) where $V = 0$? Explain.

23.49 •• A very long uniform line of charge with charge per unit length $\lambda = +5.00\ \mu\text{C/m}$ lies along the x-axis, with its midpoint at the origin. A very large uniform sheet of charge is parallel to the xy-plane; the center of the sheet is at $z = +0.600$ m. The sheet has charge per unit area $\sigma = +8.00\ \mu\text{C/m}^2$, and the center of the sheet is at $x = 0$, $y = 0$. Point A is on the z-axis at $z = +0.300$ m, and point B is on the z-axis at $z = -0.200$ m. What is the potential difference $V_{AB} = V_A - V_B$ between points A and B? Which point, A or B, is at higher potential?

23.50 ••• A small sphere with mass 5.00×10^{-7} kg and charge $+7.00\ \mu\text{C}$ is released from rest a distance of 0.400 m above a large horizontal insulating sheet of charge that has uniform surface charge density $\sigma = +8.00\ \text{pC/m}^2$. Using energy methods, calculate the speed of the sphere when it is 0.100 m above the sheet.

23.51 •• A gold nucleus has a radius of 7.3×10^{-15} m and a charge of $+79e$. Through what voltage must an alpha particle, with charge $+2e$, be accelerated so that it has just enough energy to reach a distance of 2.0×10^{-14} m from the surface of a gold nucleus? (Assume that the gold nucleus remains stationary and can be treated as a point charge.)

23.52 •• **CP** A proton and an alpha particle are released from rest when they are 0.225 nm apart. The alpha particle (a helium nucleus) has essentially four times the mass and two times the charge of a proton. Find the maximum *speed* and maximum *acceleration* of each of these particles. When do these maxima occur: just following the release of the particles or after a very long time?

23.53 • A particle with charge +7.60 nC is in a uniform electric field directed to the left. Another force, in addition to the electric force, acts on the particle so that when it is released from rest, it moves to the right. After it has moved 8.00 cm, the additional force has done 6.50×10^{-5} J of work and the particle has 4.35×10^{-5} J of kinetic energy. (a) What work was done by the electric force? (b) What is the potential of the starting point with respect to the end point? (c) What is the magnitude of the electric field?

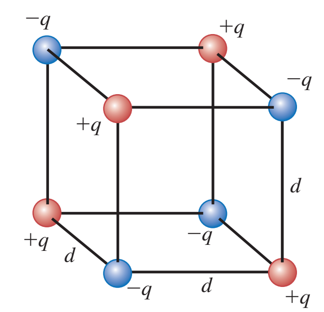
23.54 •• Identical charges $q = +5.00\ \mu\text{C}$ are placed at opposite corners of a square that has sides of length 8.00 cm. Point A is at one of the empty corners, and point B is at the center of the square. A charge $q_0 = -3.00\ \mu\text{C}$ is placed at point A and moves along the diagonal of the square to point B. (a) What is the magnitude of the net electric force on q_0 when it is at point A? Sketch the placement of the charges and the direction of the net force. (b) What is the magnitude of the net electric force on q_0 when it is at point B? (c) How much work does the electric force do on q_0 during its motion from A to B? Is this work positive or negative? When it goes from A to B, does q_0 move to higher potential or to lower potential?

23.55 •• **CALC** A vacuum tube diode consists of concentric cylindrical electrodes, the negative cathode and the positive anode. Because of the accumulation of charge near the cathode, the electric potential between the electrodes is given by $V(x) = Cx^{4/3}$ where x is the distance from the cathode and C is a constant, characteristic of a particular diode and operating conditions. Assume that the distance between the cathode and anode is 13.0 mm and the potential difference between electrodes is 240 V. (a) Determine the value of C. (b) Obtain a formula for the electric field between the electrodes as a function of x. (c) Determine the force on an electron when the electron is halfway between the electrodes.

23.56 •• When you scuff your feet on a carpet, you gain electrons and become negatively charged. If you then place your finger near a metallic surface, such as a doorknob, an electric field develops between your finger and the doorknob. As your finger gets closer to the surface, the magnitude of the electric field increases. When it exceeds a threshold of 3 MV/m, the air breaks down, creating a small "lightning" strike, which you feel as a shock. (a) Estimate the distance between your finger and a doorknob at the point you feel the shock. (b) Using this estimate and the threshold field strength of 3 MV/m, find the potential between your finger and the doorknob. (c) At this (small) distance, we can treat the tip of your finger and the end of the doorknob as infinite planar surfaces with opposite charge density. Estimate that density. (d) Estimate the effective area of your fingertip as presented to the doorknob. (e) Use these results to estimate the magnitude of charge on your finger. (f) Assuming that your net excess charge has built up on your finger, attracted there by the doorknob, estimate the number of electrons that you lost while scuffing your feet.

23.57 •• **An Ionic Crystal. Figure P23.57** shows eight point charges arranged at the corners of a cube with sides of length d. The values of the charges are $+q$ and $-q$, as shown. This is a model of one cell of a cubic ionic crystal. In sodium chloride (NaCl), for instance, the positive ions are Na^+ and the negative ions are Cl^-. (a) Calculate the potential energy U of this arrangement. (Take as zero the potential energy of the eight charges when they are infinitely far apart.) (b) In part (a), you should have found that $U < 0$. Explain the relationship between this result and the observation that such ionic crystals exist in nature.

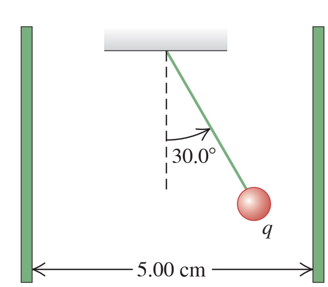
Figure **P23.57**



23.58 •• Electrical power is transmitted to our homes by overhead wires strung between poles. A typical residential utility line has a maximum potential of 22 kV relative to ground. We can treat this potential as constant and model it as generated by a net charge distributed on the wire. (a) Estimate the height of an electrical transmission line. (b) Treat the electrical wire as a long conducting cylinder with a radius of 2 cm. If the potential between the surface of the wire and a position directly beneath the wire is 22 kV, what is the linear charge density on the wire? (c) Use your estimate of the linear charge density to estimate the strength of the electric field on the ground beneath the wire.

23.59 •• **CP** A small sphere with mass 1.50 g hangs by a thread between two very large parallel vertical plates 5.00 cm apart (**Fig. P23.59**). The plates are insulating and have uniform surface charge densities $+\sigma$ and $-\sigma$. The charge on the sphere is $q = 8.90 \times 10^{-6}$ C. What potential difference between the plates will cause the thread to assume an angle of 30.0° with the vertical?

Figure **P23.59**



23.60 •• Two spherical shells have a common center. The inner shell has radius $R_1 = 5.00$ cm and charge $q_1 = +3.00 \times 10^{-6}$ C; the outer shell has radius $R_2 = 15.0$ cm and charge $q_2 = -5.00 \times 10^{-6}$ C. Both charges are spread uniformly over the shell surface. What is the electric potential due to the two shells at the following distances from their common center: (a) $r = 2.50$ cm; (b) $r = 10.0$ cm; (c) $r = 20.0$ cm? Take $V = 0$ at a large distance from the shells.

23.61 • CALC Coaxial Cylinders. A long metal cylinder with radius a is supported on an insulating stand on the axis of a long, hollow, metal tube with radius b. The positive charge per unit length on the inner cylinder is λ, and there is an equal negative charge per unit length on the outer cylinder. (a) Calculate the potential $V(r)$ for (i) $r < a$; (ii) $a < r < b$; (iii) $r > b$. (*Hint:* The net potential is the sum of the potentials due to the individual conductors.) Take $V = 0$ at $r = b$. (b) Show that the potential of the inner cylinder with respect to the outer is

$$V_{ab} = \frac{\lambda}{2\pi\epsilon_0} \ln \frac{b}{a}$$

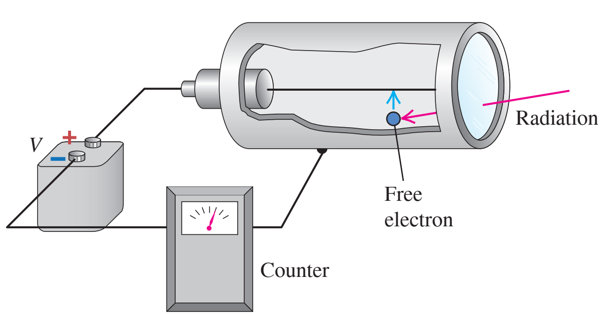
(c) Use Eq. (23.23) and the result from part (a) to show that the electric field at any point between the cylinders has magnitude

$$E(r) = \frac{V_{ab}}{\ln(b/a)} \frac{1}{r}$$

(d) What is the potential difference between the two cylinders if the outer cylinder has no net charge?

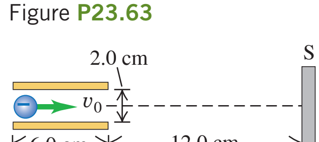
23.62 •• A *Geiger counter* detects radiation such as alpha particles by using the fact that the radiation ionizes the air along its path. A thin wire lies on the axis of a hollow metal cylinder and is insulated from it (**Fig. P23.62**). A large potential difference is established between the wire and the outer cylinder, with the wire at higher potential; this sets up a strong electric field directed radially outward. When ionizing radiation enters the device, it ionizes a few air molecules. The free electrons produced are accelerated by the electric field toward the wire and, on the way there, ionize many more air molecules. Thus a current pulse is produced that can be detected by appropriate electronic circuitry and converted to an audible "click." Suppose the radius of the central wire is 145 μm and the radius of the hollow cylinder is 1.80 cm. What potential difference between the wire and the cylinder produces an electric field of 2.00×10^4 V/m at a distance of 1.20 cm from the axis of the wire? (The wire and cylinder are both very long in comparison to their radii, so the results of Problem 23.61 apply.)

Figure P23.62



23.63 • CP Deflection in a CRT. Cathode-ray tubes (CRTs) were often found in oscilloscopes and computer monitors. In **Fig. P23.63** an electron with an initial speed of 6.50×10^6 m/s is projected along the axis midway between the deflection plates of a cathode-ray tube. The potential difference between the two plates is 22.0 V and the lower plate is the one at higher potential. (a) What is the force (magnitude and direction) on the electron when it is between the plates? (b) What is the acceleration of the electron (magnitude and direction) when acted on by the force in part (a)? (c) How far below the axis has the electron moved when it reaches the end of the plates? (d) At what angle with the axis is it moving as it leaves the plates? (e) How far below the axis will it strike the fluorescent screen S?

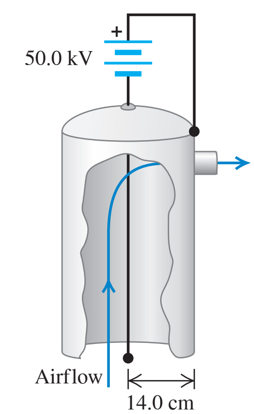
Figure P23.63



23.64 •• CP Deflecting Plates of an Oscilloscope. The vertical deflecting plates of a typical classroom oscilloscope are a pair of parallel square metal plates carrying equal but opposite charges. Typical dimensions are about 3.0 cm on a side, with a separation of about 5.0 mm. The potential difference between the plates is 25.0 V. The plates are close enough that we can ignore fringing at the ends. Under these conditions: (a) how much charge is on each plate, and (b) how strong is the electric field between the plates? (c) If an electron is ejected at rest from the negative plate, how fast is it moving when it reaches the positive plate?

23.65 •• *Electrostatic precipitators* use electric forces to remove pollutant particles from smoke, in particular in the smokestacks of coal-burning power plants. One form of precipitator consists of a vertical, hollow, metal cylinder with a thin wire, insulated from the cylinder, running along its axis (**Fig. P23.65**). A large potential difference is established between the wire and the outer cylinder, with the wire at lower potential. This sets up a strong radial electric field directed inward. The field produces a region of ionized air near the wire. Smoke enters the precipitator at the bottom, ash and dust in it pick up electrons, and the charged pollutants are accelerated toward the outer cylinder wall by the electric field. Suppose the radius of the central wire is 90.0 μm, the radius of the cylinder is 14.0 cm, and a potential difference of 50.0 kV is established between the wire and the cylinder. Also assume that the wire and cylinder are both very long in comparison to the cylinder radius, so the results of Problem 23.61 apply. (a) What is the magnitude of the electric field midway between the wire and the cylinder wall? (b) What magnitude of charge must a 30.0 μg ash particle have if the electric field computed in part (a) is to exert a force ten times the weight of the particle?

Figure P23.65



23.66 •• CALC A disk with radius R has uniform surface charge density σ. (a) By regarding the disk as a series of thin concentric rings, calculate the electric potential V at a point on the disk's axis a distance x from the center of the disk. Assume that the potential is zero at infinity. (*Hint:* Use the result of Example 23.11 in Section 23.3.) (b) Calculate $-\partial V/\partial x$. Show that the result agrees with the expression for E_x calculated in Example 21.11 (Section 21.5).

23.67 ••• CALC Self-Energy of a Sphere of Charge. A solid sphere of radius R contains a total charge Q distributed uniformly throughout its volume. Find the energy needed to assemble this charge by bringing infinitesimal charges from far away. This energy is called the "self-energy" of the charge distribution. (*Hint:* After you have assembled a charge q in a sphere of radius r, how much energy would it take to add a spherical shell of thickness dr having charge dq? Then integrate to get the total energy.)

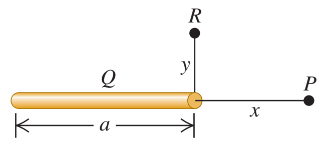
23.68 • CALC A thin insulating rod is bent into a semicircular arc of radius a, and a total electric charge Q is distributed uniformly along the rod. Calculate the potential at the center of curvature of the arc if the potential is assumed to be zero at infinity.

23.69 •• Charge $Q = +4.00\ \mu$C is distributed uniformly over the volume of an insulating sphere that has radius $R = 5.00$ cm. What is the potential difference between the center of the sphere and the surface of the sphere?

23.70 • (a) If a spherical raindrop of radius 0.650 mm carries a charge of -3.60 pC uniformly distributed over its volume, what is the potential at its surface? (Take the potential to be zero at an infinite distance from the raindrop.) (b) Two identical raindrops, each with radius and charge specified in part (a), collide and merge into one larger raindrop. What is the radius of this larger drop, and what is the potential at its surface, if its charge is uniformly distributed over its volume?

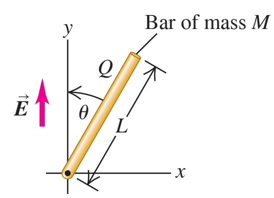
23.71 • **CALC** Electric charge is distributed uniformly along a thin rod of length a, with total charge Q. Take the potential to be zero at infinity. Find the potential at the following points (**Fig. P23.71**): (a) point P, a distance x to the right of the rod, and (b) point R, a distance y above the right-hand end of the rod. (c) In parts (a) and (b), what does your result reduce to as x or y becomes much larger than a?

Figure **P23.71**



23.72 ••• **CP CALC** A rigid bar with mass M, length L, and a uniformly distributed positive charge Q is free to pivot about the origin in the presence of a spatially uniform electric field $\vec{E} = E\hat{\jmath}$, as shown in **Fig. P23.72**. Assume that the electric forces are large enough for gravity to be neglected. (a) Write the potential $V(y)$ due to the electric field as a function of y, using the convention $V(0) = V_0$, where V_0 is a constant to be determined. (b) Determine the electric potential energy U of the system as a function of θ and V_0. (*Hint:* $U = \int V\,dq$.) (c) For what value of V_0 does the potential energy vanish when $\theta = 0$? (d) If the bar is released from rest at the position $\theta = 90°$, what is the angular speed of the bar as it passes the position $\theta = 0$? (e) If the bar starts from rest at a small value of θ, with what frequency will the bar oscillate around the position $\theta = 0$?

Figure **P23.72**



23.73 ••• **CP** A helium nucleus, also known as an α (alpha) particle, consists of two protons and two neutrons and has a diameter of 10^{-15} m = 1 fm. The protons, with a charge of $+e$, are subject to a repulsive Coulomb force. Since the neutrons have zero charge, there must be an attractive force that counteracts the electric repulsion and keeps the protons from flying apart. This so-called strong force plays a central role in particle physics. (a) As a crude model, assume that an α particle consists of two pointlike protons attracted by a Hooke's-law spring with spring constant k, and ignore the neutrons. Assume further that in the absence of other forces, the spring has an equilibrium separation of zero. Write an expression for the potential energy when the protons are separated by distance d. (b) Minimize this potential to find the equilibrium separation d_0 in terms of e and k. (c) If $d_0 = 1.00$ fm, what is the value of k? (d) How much energy is stored in this system, in terms of electron volts? (e) A proton has a mass of 1.67×10^{-27} kg. If the spring were to break, the α particle would disintegrate and the protons would fly off in opposite directions. What would be their ultimate speed?

23.74 • A metal sphere with radius R_1 has a charge Q_1. Take the electric potential to be zero at an infinite distance from the sphere. (a) What are the electric field and electric potential at the surface of the sphere? This sphere is now connected by a long, thin conducting wire to another sphere of radius R_2 that is several meters from the first sphere. Before the connection is made, this second sphere is uncharged. After electrostatic equilibrium has been reached, what are (b) the total charge on each sphere; (c) the electric potential at the surface of each sphere; (d) the electric field at the surface of each sphere? Assume that the amount of charge on the wire is much less than the charge on each sphere.

23.75 • An alpha particle with kinetic energy 9.50 MeV (when far away) collides head-on with a lead nucleus at rest. What is the distance of closest approach of the two particles? (Assume that the lead nucleus remains stationary and may be treated as a point charge. The atomic number of lead is 82. The alpha particle is a helium nucleus, with atomic number 2.)

23.76 • **CALC** The electric potential V in a region of space is given by

$$V(x, y, z) = A(x^2 - 3y^2 + z^2)$$

where A is a constant. (a) Derive an expression for the electric field $\vec{E}$ at any point in this region. (b) The work done by the field when a 1.50 μC test charge moves from the point $(x, y, z) = (0, 0, 0.250\text{ m})$ to the origin is measured to be 6.00×10^{-5} J. Determine A. (c) Determine the electric field at the point (0, 0, 0.250 m). (d) Show that in every plane parallel to the xz-plane the equipotential contours are circles. (e) What is the radius of the equipotential contour corresponding to $V = 1280$ V and $y = 2.00$ m?

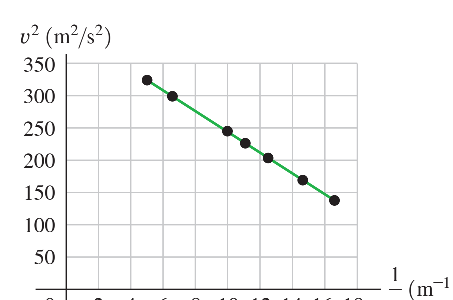
23.77 •• **DATA** The electric potential in a region that is within 2.00 m of the origin of a rectangular coordinate system is given by $V = Ax^l + By^m + Cz^n + D$, where $A, B, C, D, l, m,$ and n are constants. The units of $A, B, C,$ and D are such that if $x, y,$ and z are in meters, then V is in volts. You measure V and each component of the electric field at four points and obtain these results:

Point	(x, y, z) (m)	V (V)	E_x (V/m)	E_y (V/m)	E_z (V/m)
1	(0, 0, 0)	10.0	0	0	0
2	(1.00, 0, 0)	4.0	12.0	0	0
3	(0, 1.00, 0)	6.0	0	12.0	0
4	(0, 0, 1.00)	8.0	0	0	12.0

(a) Use the data in the table to calculate $A, B, C, D, l, m,$ and n. (b) What are V and the magnitude of E at the points (0, 0, 0), (0.50 m, 0.50 m, 0.50 m), and (1.00 m, 1.00 m, 1.00 m)?

23.78 •• **DATA** A small, stationary sphere carries a net charge Q. You perform the following experiment to measure Q: From a large distance you fire a small particle with mass $m = 4.00 \times 10^{-4}$ kg and charge $q = 5.00 \times 10^{-8}$ C directly at the center of the sphere. The apparatus you are using measures the particle's speed v as a function of the distance x from the sphere. The sphere's mass is much greater than the mass of the projectile particle, so you assume that the sphere remains at rest. All of the measured values of x are much larger than the radius of either object, so you treat both objects as point particles. You plot your data on a graph of v^2 versus $(1/x)$ (**Fig. P23.78**). The straight line $v^2 = 400\text{ m}^2/\text{s}^2 - [(15.75\text{ m}^3/\text{s}^2)/x]$ gives a good fit to the data points. (a) Explain why the graph is a straight line. (b) What is the initial speed v_0 of the particle when it is very far from the sphere? (c) What is Q? (d) How close does the particle get to the sphere? Assume that this distance is much larger than the radii of the particle and sphere, so continue to treat them as point particles and to assume that the sphere remains at rest.

Figure **P23.78**



23.79 ••• DATA The Millikan Oil-Drop Experiment. The charge of an electron was first measured by the American physicist Robert Millikan during 1909–1913. In his experiment, oil was sprayed in very fine drops (about 10^{-4} mm in diameter) into the space between two parallel horizontal plates separated by a distance d. A potential difference V_{AB} was maintained between the plates, causing a downward electric field between them. Some of the oil drops acquired a negative charge because of frictional effects or because of ionization of the surrounding air by x rays or radioactivity. The drops were observed through a microscope. (a) Show that an oil drop of radius r at rest between the plates remained at rest if the magnitude of its charge was

$$q = \frac{4\pi}{3}\frac{\rho r^3 g d}{V_{AB}}$$

where ρ is oil's density. (Ignore the buoyant force of the air.) By adjusting V_{AB} to keep a given drop at rest, Millikan determined the charge on that drop, provided its radius r was known. (b) Millikan's oil drops were much too small to measure their radii directly. Instead, Millikan determined r by cutting off the electric field and measuring the *terminal speed* v_t of the drop as it fell. (We discussed terminal speed in Section 5.3.) The viscous force F on a sphere of radius r moving at speed v through a fluid with viscosity η is given by Stokes's law: $F = 6\pi\eta r v$. When a drop fell at v_t, the viscous force just balanced the drop's weight $w = mg$. Show that the magnitude of the charge on the drop was

$$q = 18\pi\frac{d}{V_{AB}}\sqrt{\frac{\eta^3 v_t^3}{2\rho g}}$$

(c) You repeat the Millikan oil-drop experiment. Four of your measured values of V_{AB} and v_t are listed in the table:

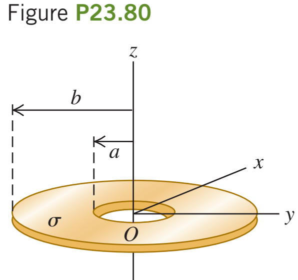
Drop	1	2	3	4
V_{AB} (V)	9.16	4.57	12.32	6.28
v_t (10^{-5} m/s)	2.54	0.767	4.39	1.52

In your apparatus, the separation d between the horizontal plates is 1.00 mm. The density of the oil you use is 824 kg/m^3. For the viscosity η of air, use the value 1.81×10^{-5} N·s/m^2. Assume that $g = 9.80$ m/s^2. Calculate the charge q of each drop. (d) If electric charge is *quantized* (that is, exists in multiples of the magnitude of the charge of an electron), then the charge on each drop is $-ne$, where n is the number of excess electrons on each drop. (All four drops in your table have negative charge.) Drop 2 has the smallest magnitude of charge observed in the experiment, for all 300 drops on which measurements were made, so assume that its charge is due to an excess charge of one electron. Determine the number of excess electrons n for each of the other three drops. (e) Use $q = -ne$ to calculate e from the data for each of the four drops, and average these four values to get your best experimental value of e.

CHALLENGE PROBLEMS

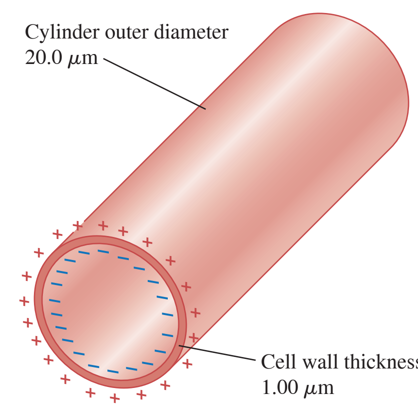
23.80 ••• CALC An annulus with an inner radius of a and an outer radius of b has charge density σ and lies in the xy-plane with its center at the origin, as shown in **Fig. P23.80**. (a) Using the convention that the potential vanishes at infinity, determine the potential at all points on the z-axis. (b) Determine the electric field at all points on the z-axis by differentiating the potential. (c) Show that in the limit $a \to 0$, $b \to \infty$ the electric field reproduces the result obtained in Example 22.7 for an infinite plane sheet of charge. (d) If $a = 5.00$ cm, $b = 10.0$ cm and the total charge on the annulus is 1.00 μC, what is the potential at the origin? (e) If a particle with mass 1.00 g (much less than the mass of the annulus) and charge 1.00 μC is placed at the origin and given the slightest nudge, it will be projected along the z-axis. In this case, what will be its ultimate speed?

Figure **P23.80**



23.81 ••• BIO A heart cell can be modeled as a cylindrical shell that is 100 μm long, with an outer diameter of 20.0 μm and a cell wall thickness of 1.00 μm, as shown in **Fig. P23.81**. Potassium ions move across the cell wall, depositing positive charge on the outer surface and leaving a net negative charge on the inner surface. During the so-called resting phase, the inside of the cell has a potential that is 90.0 mV lower than the potential on its outer surface. (a) If the net charge of the cell is zero, what is the magnitude of the total charge on either cell wall membrane? Ignore edge effects and treat the cell as a very long cylinder. (b) What is the magnitude of the electric field just inside the cell wall? (c) In a subsequent depolarization event, sodium ions move through channels in the cell wall, so that the inner membrane becomes positively charged. At the end of this event, the inside of the cell has a potential that is 20.0 mV higher than the potential outside the cell. If we model this event by charge moving from the outer membrane to the inner membrane, what magnitude of charge moves across the cell wall during this event? (d) If this were done entirely by the motion of sodium ions, Na^+, how many ions have moved?

Figure **P23.81**



23.82 ••• CALC A hollow, thin-walled insulating cylinder of radius R and length L (like the cardboard tube in a roll of toilet paper) has charge Q uniformly distributed over its surface. (a) Calculate the electric potential at all points along the axis of the tube. Take the origin to be at the center of the tube, and take the potential to be zero at infinity. (b) Show that if $L \ll R$, the result of part (a) reduces to the potential on the axis of a ring of charge of radius R. (See Example 23.11 in Section 23.3.) (c) Use the result of part (a) to find the electric field at all points along the axis of the tube.

23.83 ••• CP In experiments in which atomic nuclei collide, head-on collisions like that described in Problem 23.75 do happen, but "near misses" are more common. Suppose the alpha particle in that problem is not "aimed" at the center of the lead nucleus but has an initial nonzero angular momentum (with respect to the stationary lead nucleus) of magnitude $L = p_0 b$, where p_0 is the magnitude of the particle's initial momentum and $b = 1.00 \times 10^{-12}$ m. What is the distance of closest approach? Repeat for $b = 1.00 \times 10^{-13}$ m and $b = 1.00 \times 10^{-14}$ m.

MCAT-STYLE PASSAGE PROBLEMS

Materials Analysis with Ions. *Rutherford backscattering spectrometry* (RBS) is a technique used to determine the structure and composition of materials. A beam of ions (typically helium ions) is accelerated to high energy and aimed at a sample. By analyzing the distribution and energy of the ions that are scattered from (that is, deflected by collisions with) the atoms in the sample, researchers can determine the sample's composition. To accelerate the ions to high energies, a *tandem electrostatic accelerator* may be used. In this device, negative ions (He^-) start at a potential $V = 0$ and are accelerated by a high positive voltage at the midpoint of the accelerator. The high voltage produces a constant electric field in the acceleration tube through which the ions move. When accelerated ions reach the midpoint, the electrons are stripped off, turning the negative ions into doubly positively charged ions (He^{++}). These positive ions are then repelled from the midpoint by the high positive voltage there and continue to accelerate to the far end of the accelerator, where again $V = 0$.

23.84 For a particular experiment, helium ions are to be given a kinetic energy of 3.0 MeV. What should the voltage at the center of the accelerator be, assuming that the ions start essentially at rest? (a) −3.0 MV; (b) +3.0 MV; (c) +1.5 MV; (d) +1.0 MV.

23.85 A helium ion (He^{++}) that comes within about 10 fm of the center of the nucleus of an atom in the sample may induce a nuclear reaction instead of simply scattering. Imagine a helium ion with a kinetic energy of 3.0 MeV heading straight toward an atom at rest in the sample. Assume that the atom stays fixed. What minimum charge can the nucleus of the atom have such that the helium ion gets no closer than 10 fm from the center of the atomic nucleus? ($1 \text{ fm} = 1 \times 10^{-15}$ m, and e is the magnitude of the charge of an electron or a proton.) (a) $2e$; (b) $11e$; (c) $20e$; (d) $22e$.

23.86 The maximum voltage at the center of a typical tandem electrostatic accelerator is 6.0 MV. If the distance from one end of the acceleration tube to the midpoint is 12 m, what is the magnitude of the average electric field in the tube under these conditions? (a) 41,000 V/m; (b) 250,000 V/m; (c) 500,000 V/m; (d) 6,000,000 V/m.

ANSWERS

Chapter Opening Question ?

(iii) A large, constant potential difference V_{ab} is maintained between the welding tool (a) and the metal pieces to be welded (b). For a given potential difference between two conductors a and b, the smaller the distance d separating the conductors, the greater is the magnitude E of the field between them. Hence d must be small in order for E to be large enough to ionize the gas between the conductors (see Section 23.3) and produce an arc through this gas.

Key Example VARIATION Problems

VP23.2.1 (a) 1.51×10^{-16} J (b) 2.32×10^{-15} J
VP23.2.2 (a) 5.22×10^{-15} J (b) 1.92×10^{-18} C
VP23.2.3 (a) $-e^2/24\pi\epsilon_0 a$ (b) $-7e^2/24\pi\epsilon_0 a$
VP23.2.4 (a) 7.21×10^{-9} C (b) 5.59×10^{-6} J
VP23.7.1 (a) +25.0 V (b) +15.0 V (c) 0 (d) −25.0 V
VP23.7.2 (a) 3.6×10^3 V (b) 2.2×10^3 V (c) 1.6×10^3 V (d) point (a); no
VP23.7.3 (a) +133 V (b) −133 V (c) 1.98×10^5 m/s
VP23.7.4 (a) $Q/8\pi\epsilon_0 R$ (b) the center
VP23.12.1 (a) $qq_0/30\pi\epsilon_0 R$ (b) $-5qq_0/56\pi\epsilon_0 R$ (c) $3qq_0/16\pi\epsilon_0 R$
VP23.12.2 (a) +16.0 V (b) 3.20×10^{-8} J (c) lower, 3.58 m/s
VP23.12.3 (a) 1.61×10^{-6} J (b) 5.39×10^{-6} J (c) $K = 3.42 \times 10^{-6}$ J, $v = 41.3$ m/s
VP23.12.4 (a) $dq = (Q/L)dx$ (b) $(Q/4\pi\epsilon_0 L)\ln 2$

Bridging Problem

$$\frac{qQ}{8\pi\epsilon_0 a}\ln\left(\frac{L+a}{L-a}\right)$$

? In flash photography, the energy used to make the flash is stored in a capacitor, which consists of two closely spaced conductors that carry opposite charges. If the amount of charge on the conductors is doubled, by what factor does the stored energy increase? (i) $\sqrt{2}$; (ii) 2; (iii) $2\sqrt{2}$; (iv) 4; (v) 8.

24 Capacitance and Dielectrics

When you stretch the rubber band of a slingshot or pull back the string of an archer's bow, you are storing mechanical energy as elastic potential energy. A capacitor is a device that stores *electric* potential energy and electric charge. To make a capacitor, just insulate two conductors from each other. To store energy in this device, transfer charge from one conductor to the other so that one has a negative charge and the other has an equal amount of positive charge. Work must be done to move the charges through the resulting potential difference between the conductors, and the work done is stored as electric potential energy.

Capacitors have a tremendous number of practical applications in devices such as electronic flash units for photography, mobile phones, airbag sensors for cars, and radio and television receivers. We'll encounter many of these applications in later chapters (particularly Chapter 31, in which we'll see the crucial role played by capacitors in the alternating-current circuits that pervade our technological society). In this chapter, however, our emphasis is on the fundamental properties of capacitors. For a particular capacitor, the ratio of the charge on each conductor to the potential difference between the conductors is a constant, called the *capacitance*. The capacitance depends on the sizes and shapes of the conductors and on the insulating material (if any) between them. Compared to the case in which there is only vacuum between the conductors, the capacitance increases when an insulating material (a *dielectric*) is present. This happens because a redistribution of charge, called *polarization,* takes place within the insulating material. Studying polarization will give us added insight into the electrical properties of matter.

Capacitors also give us a new way to think about electric potential energy. The energy stored in a charged capacitor is related to the electric field in the space between the conductors. We'll see that electric potential energy can be regarded as being stored *in the field itself.* The idea that the electric field is itself a storehouse of energy is at the heart of the theory of electromagnetic waves and our modern understanding of the nature of light, to be discussed in Chapter 32.

LEARNING OUTCOMES

In this chapter, you'll learn...

24.1 The nature of capacitors, and how to calculate a quantity that measures their ability to store charge.

24.2 How to analyze capacitors connected in a network.

24.3 How to calculate the amount of energy stored in a capacitor.

24.4 What dielectrics are, and how they make capacitors more effective.

24.5 How a dielectric inside a charged capacitor becomes polarized.

24.6 How to use Gauss's laws when dielectrics are present.

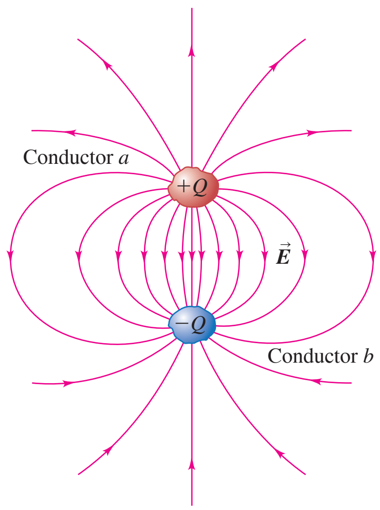
You'll need to review...

21.2, 21.5, 21.7 Polarization; field of charged conductors; electric dipoles.

22.3–22.5 Gauss's law.

23.3, 23.4 Potential for charged conductors; potential due to a cylindrical charge distribution.

Figure **24.1** Any two conductors a and b insulated from each other form a capacitor.



24.1 CAPACITORS AND CAPACITANCE

Any two conductors separated by an insulator (or a vacuum) form a **capacitor** (**Fig. 24.1**). In most practical applications, each conductor initially has zero net charge and electrons are transferred from one conductor to the other; this is called *charging* the capacitor. Then the two conductors have charges with equal magnitude and opposite sign, and the *net* charge on the capacitor as a whole remains zero. We'll assume throughout this chapter that this is the case. When we say that a capacitor has charge Q, or that a charge Q is *stored* on the capacitor, we mean that the conductor at higher potential has charge $+Q$ and the conductor at lower potential has charge $-Q$ (assuming that Q is positive). Keep this in mind in the following discussion and examples.

In circuit diagrams a capacitor is represented by either of these symbols:

The vertical lines (straight or curved) represent the conductors, and the horizontal lines represent wires connected to either conductor. One common way to charge a capacitor is to connect these two wires to opposite terminals of a battery. Once the charges Q and $-Q$ are established on the conductors, the battery is disconnected. This gives a fixed *potential difference* V_{ab} between the conductors (that is, the potential of the positively charged conductor a with respect to the negatively charged conductor b) that is just equal to the voltage of the battery.

The electric field at any point in the region between the conductors is proportional to the magnitude Q of charge on each conductor. It follows that the potential difference V_{ab} between the conductors is also proportional to Q. If we double the magnitude of charge on each conductor, the charge density at each point doubles, the electric field at each point doubles, and the potential difference between conductors doubles; however, the *ratio* of charge to potential difference does not change. This ratio is called the **capacitance** C of the capacitor:

Capacitance of a capacitor

$$C = \frac{Q}{V_{ab}} \qquad (24.1)$$

Q: Magnitude of charge on each conductor

V_{ab}: Potential difference between conductors (a has charge $+Q$, b has charge $-Q$)

The SI unit of capacitance is called one **farad** (1 F), in honor of the 19th-century English physicist Michael Faraday. From Eq. (24.1), one farad is equal to one *coulomb per volt* (1 C/V):

$$1\ \text{F} = 1\ \text{farad} = 1\ \text{C/V} = 1\ \text{coulomb/volt}$$

CAUTION **Capacitance vs. coulombs** Don't confuse the symbol C for capacitance (which is always in italics) with the abbreviation C for coulombs (which is never italicized).

The greater the capacitance C of a capacitor, the greater the magnitude Q of charge on either conductor for a given potential difference V_{ab} and hence the greater the amount of stored energy. (Remember that potential is potential energy per unit charge.) Thus *capacitance is a measure of the ability of a capacitor to store energy.* We'll see that the capacitance depends only on the shapes, sizes, and relative positions of the conductors and on the nature of the insulator between them. (For special types of insulating materials, the capacitance *does* depend on Q and V_{ab}. We won't discuss these materials, however.)

Calculating Capacitance: Capacitors in Vacuum

We can calculate the capacitance C of a given capacitor by finding the potential difference V_{ab} between the conductors for a given magnitude of charge Q and then using Eq. (24.1). For now we'll consider only *capacitors in vacuum;* that is, empty space separates the conductors that make up the capacitor.

The simplest form of capacitor consists of two parallel conducting plates, each with area A, separated by a distance d that is small in comparison with their dimensions (**Fig. 24.2a**). When the plates are charged, the electric field is almost completely localized in the region between the plates (Fig. 24.2b). As we discussed in Example 22.8 (Section 22.4), the field between such plates is essentially *uniform,* and the charges on the plates are uniformly distributed over their opposing surfaces. We call this arrangement a **parallel-plate capacitor.**

We found the electric-field magnitude E for this arrangement in Example 21.12 (Section 21.5) by using the principle of superposition of electric fields and again in Example 22.8 (Section 22.4) by using Gauss's law. It would be a good idea to review those examples. We found that $E = \sigma/\epsilon_0$, where σ is the magnitude (absolute value) of the surface charge density on each plate. This is equal to the magnitude of the total charge Q on each plate divided by the area A of the plate, or $\sigma = Q/A$, so the field magnitude E can be expressed as

$$E = \frac{\sigma}{\epsilon_0} = \frac{Q}{\epsilon_0 A}$$

The field is uniform and the distance between the plates is d, so the potential difference (voltage) between the two plates is

$$V_{ab} = Ed = \frac{1}{\epsilon_0}\frac{Qd}{A}$$

Thus

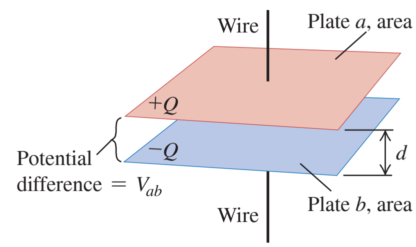
Capacitance of a parallel-plate capacitor in vacuum

$$C = \frac{Q}{V_{ab}} = \epsilon_0 \frac{A}{d} \qquad (24.2)$$

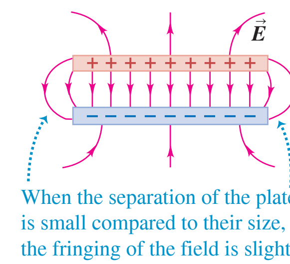
Magnitude of charge on each plate: Q
Potential difference between plates: V_{ab}
Area of each plate: A
Distance between plates: d
Electric constant: ϵ_0

Figure **24.2** A charged parallel-plate capacitor.

(a) Arrangement of the capacitor plates



(b) Side view of the electric field $\vec{E}$



The capacitance depends on only the geometry of the capacitor; it is directly proportional to the area A of each plate and inversely proportional to their separation d. The quantities A and d are constants for a given capacitor, and ϵ_0 is a universal constant. Thus in vacuum the capacitance C is a constant independent of the charge on the capacitor or the potential difference between the plates. If one of the capacitor plates is flexible, C changes as the plate separation d changes. This is the operating principle of a condenser microphone (**Fig. 24.3**).

When matter is present between the plates, its properties affect the capacitance. We'll return to this topic in Section 24.4. Meanwhile, we remark that if the space contains air at atmospheric pressure instead of vacuum, the capacitance differs from the prediction of Eq. (24.2) by less than 0.06%.

In Eq. (24.2), if A is in square meters and d in meters, then C is in farads. The units of the electric constant ϵ_0 are $\mathrm{C^2/N \cdot m^2}$, so

$$1\ \mathrm{F} = 1\ \mathrm{C^2/N \cdot m} = 1\ \mathrm{C^2/J}$$

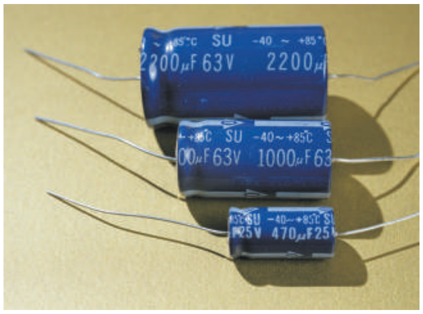
Because $1\ \mathrm{V} = 1\ \mathrm{J/C}$ (energy per unit charge), this equivalence agrees with our definition $1\ \mathrm{F} = 1\ \mathrm{C/V}$. Finally, we can express the units of ϵ_0 as $1\ \mathrm{C^2/N \cdot m^2} = 1\ \mathrm{F/m}$, so

$$\epsilon_0 = 8.85 \times 10^{-12}\ \mathrm{F/m}$$

This relationship is useful in capacitance calculations, and it also helps us to verify that Eq. (24.2) is dimensionally consistent.

Figure **24.3** Inside a condenser microphone is a capacitor with one rigid plate and one flexible plate. The two plates are kept at a constant potential difference V_{ab}. Sound waves cause the flexible plate to move back and forth, varying the capacitance C and causing charge to flow to and from the capacitor in accordance with the relationship $C = Q/V_{ab}$. Thus a sound wave is converted to a charge flow that can be amplified and recorded digitally.

Figure **24.4** A commercial capacitor is labeled with the value of its capacitance. For these capacitors, $C = 2200\ \mu\text{F}$, $1000\ \mu\text{F}$, and $470\ \mu\text{F}$.



One farad is a very large capacitance, as the following example shows. In many applications the most convenient units of capacitance are the *microfarad* ($1\ \mu\text{F} = 10^{-6}\ \text{F}$) and the *picofarad* ($1\ \text{pF} = 10^{-12}\ \text{F}$). For example, the flash unit in a point-and-shoot camera uses a capacitor of a few hundred microfarads (**Fig. 24.4**), while capacitances in a radio tuning circuit are typically from 10 to 100 picofarads.

For *any* capacitor in vacuum, the capacitance C depends only on the shapes, dimensions, and separation of the conductors that make up the capacitor. If the conductor shapes are more complex than those of the parallel-plate capacitor, the expression for capacitance is more complicated than in Eq. (24.2). In the following examples we show how to calculate C for two other conductor geometries.

CAUTION **Capacitor charge and potential difference** Saying that a capacitor carries charge Q means that one conductor has positive charge $+Q$ and the other conductor has negative charge $-Q$. The *net* charge on the capacitor is zero! Furthermore, the positively charged conductor is always at higher potential than the negatively charged conductor. If the potential difference V_{ab} between the conductors is 500 V, the potential of the positive conductor is 500 V greater than that of the negative conductor.

EXAMPLE 24.1 Size of a 1 F capacitor

The parallel plates of a 1.0 F capacitor are 1.0 mm apart. What is their area?

IDENTIFY and SET UP This problem uses the relationship among the capacitance C, plate separation d, and plate area A (our target variable) for a parallel-plate capacitor. We solve Eq. (24.2) for A.

EXECUTE From Eq. (24.2),

$$A = \frac{Cd}{\epsilon_0} = \frac{(1.0\ \text{F})(1.0 \times 10^{-3}\ \text{m})}{8.85 \times 10^{-12}\ \text{F/m}} = 1.1 \times 10^8\ \text{m}^2$$

EVALUATE This corresponds to a square about 10 km (about 6 miles) on a side! The volume of such a capacitor would be at least $Ad = 1.1 \times 10^5\ \text{m}^3$, equivalent to that of a cube about 50 m on a side. In fact, it's possible to make 1 F capacitors a few *centimeters* on a side. The trick is to have an appropriate substance between the plates rather than a vacuum, so that (among other things) the plate separation d can be greatly reduced. We'll explore this further in Section 24.4.

KEYCONCEPT The capacitance of a capacitor is a measure of how much charge its plates will hold for a given potential difference (voltage) between the plates. Capacitance is proportional to the area of the plates and inversely proportional to the distance between them.

EXAMPLE 24.2 Properties of a parallel-plate capacitor

WITH VARIATION PROBLEMS

The plates of a parallel-plate capacitor in vacuum are 5.00 mm apart and $2.00\ \text{m}^2$ in area. A 10.0 kV potential difference is applied across the capacitor. Compute (a) the capacitance; (b) the charge on each plate; and (c) the magnitude of the electric field between the plates.

IDENTIFY and SET UP We are given the plate area A, the plate spacing d, and the potential difference $V_{ab} = 1.00 \times 10^4\ \text{V}$ for this parallel-plate capacitor. Our target variables are the capacitance C, the charge Q on each plate, and the electric-field magnitude E. We use Eq. (24.2) to calculate C and then use Eq. (24.1) and V_{ab} to find Q. We use $E = Q/\epsilon_0 A$ to find E.

EXECUTE (a) From Eq. (24.2),

$$C = \epsilon_0 \frac{A}{d} = (8.85 \times 10^{-12}\ \text{F/m})\frac{(2.00\ \text{m}^2)}{5.00 \times 10^{-3}\ \text{m}}$$
$$= 3.54 \times 10^{-9}\ \text{F} = 0.00354\ \mu\text{F}$$

(b) The charge on the capacitor is

$$Q = CV_{ab} = (3.54 \times 10^{-9}\ \text{C/V})(1.00 \times 10^4\ \text{V})$$
$$= 3.54 \times 10^{-5}\ \text{C} = 35.4\ \mu\text{C}$$

The plate at higher potential has charge $+35.4\ \mu\text{C}$, and the other plate has charge $-35.4\ \mu\text{C}$.

(c) The electric-field magnitude is

$$E = \frac{\sigma}{\epsilon_0} = \frac{Q}{\epsilon_0 A} = \frac{3.54 \times 10^{-5}\ \text{C}}{(8.85 \times 10^{-12}\ \text{C}^2/\text{N} \cdot \text{m}^2)(2.00\ \text{m}^2)}$$
$$= 2.00 \times 10^6\ \text{N/C}$$

EVALUATE We can also find E by recalling that the electric field is equal in magnitude to the potential gradient [Eq. (23.22)]. The field between the plates is uniform, so

$$E = \frac{V_{ab}}{d} = \frac{1.00 \times 10^4\ \text{V}}{5.00 \times 10^{-3}\ \text{m}} = 2.00 \times 10^6\ \text{V/m}$$

(Remember that $1\ \text{N/C} = 1\ \text{V/m}$.)

KEYCONCEPT The electric field between the plates of a parallel-plate capacitor is uniform. To find its magnitude E, divide the potential difference between the plates by their separation; E is also proportional to the ratio of the charge on the capacitor plates to the plate's area.

EXAMPLE 24.3 A spherical capacitor

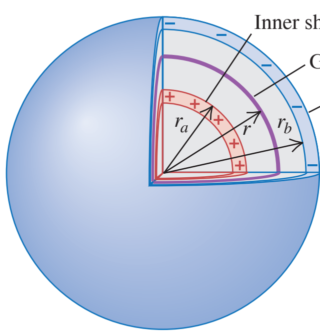
WITH VARIATION PROBLEMS

Two concentric spherical conducting shells are separated by vacuum (**Fig. 24.5**). The inner shell has total charge $+Q$ and outer radius r_a, and the outer shell has charge $-Q$ and inner radius r_b. Find the capacitance of this spherical capacitor.

IDENTIFY and SET UP By definition, the capacitance C is the magnitude Q of the charge on either sphere divided by the potential difference V_{ab} between the spheres. We first find V_{ab}, and then use Eq. (24.1) to find the capacitance $C = Q/V_{ab}$.

EXECUTE Using a Gaussian surface such as that shown in Fig. 24.5, we found in Example 22.5 (Section 22.4) that the charge on a conducting sphere produces zero field *inside* the sphere, so the outer sphere makes no contribution to the field between the spheres. Therefore the electric field *and* the electric potential between the shells are the same as those outside a charged conducting sphere with charge $+Q$. We considered that problem in Example 23.8 (Section 23.3), so the same result applies here: The potential at any point between the spheres is $V = Q/4\pi\epsilon_0 r$.

Figure **24.5** A spherical capacitor.



Hence the potential of the inner (positive) conductor at $r = r_a$ with respect to that of the outer (negative) conductor at $r = r_b$ is

$$V_{ab} = V_a - V_b = \frac{Q}{4\pi\epsilon_0 r_a} - \frac{Q}{4\pi\epsilon_0 r_b}$$

$$= \frac{Q}{4\pi\epsilon_0}\left(\frac{1}{r_a} - \frac{1}{r_b}\right) = \frac{Q}{4\pi\epsilon_0}\frac{r_b - r_a}{r_a r_b}$$

The capacitance is then

$$C = \frac{Q}{V_{ab}} = 4\pi\epsilon_0 \frac{r_a r_b}{r_b - r_a}$$

As an example, if $r_a = 9.5$ cm and $r_b = 10.5$ cm,

$$C = 4\pi(8.85 \times 10^{-12}\ \text{F/m})\frac{(0.095\ \text{m})(0.105\ \text{m})}{0.010\ \text{m}}$$

$$= 1.1 \times 10^{-10}\ \text{F} = 110\ \text{pF}$$

EVALUATE We can relate our expression for C to that for a parallel-plate capacitor. The quantity $4\pi r_a r_b$ is intermediate between the areas $4\pi r_a^2$ and $4\pi r_b^2$ of the two spheres; in fact, it's the *geometric mean* of these two areas, which we can denote by A_{gm}. The distance between spheres is $d = r_b - r_a$, so we can write $C = 4\pi\epsilon_0 r_a r_b/(r_b - r_a) = \epsilon_0 A_{gm}/d$. This has the same form as for parallel plates: $C = \epsilon_0 A/d$. If the distance between spheres is very small in comparison to their radii, their capacitance is the same as that of parallel plates with the same area and spacing.

KEYCONCEPT To find the capacitance of any capacitor made up of two conductors, first calculate the potential difference V_{ab} between the two conductors when they carry charges $+Q$ and $-Q$. The capacitance is equal to Q divided by V_{ab}.

EXAMPLE 24.4 A cylindrical capacitor

WITH VARIATION PROBLEMS

Two long, coaxial cylindrical conductors are separated by vacuum (**Fig. 24.6**). The inner cylinder has outer radius r_a and linear charge density $+\lambda$. The outer cylinder has inner radius r_b and linear charge density $-\lambda$. Find the capacitance per unit length for this capacitor.

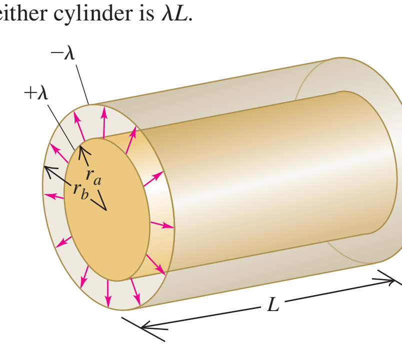
IDENTIFY and SET UP As in Example 24.3, we use the definition of capacitance, $C = Q/V_{ab}$. We use the result of Example 23.10 (Section 23.3) to find the potential difference V_{ab} between the cylinders, and find the charge Q on a length L of the cylinders from the linear charge density. We then find the corresponding capacitance C from Eq. (24.1). Our target variable is this capacitance divided by L.

EXECUTE As in Example 24.3, the potential V between the cylinders is not affected by the presence of the charged outer cylinder. Hence our result in Example 23.10 for the potential outside a charged conducting cylinder also holds in this example for potential in the space between the cylinders:

$$V = \frac{\lambda}{2\pi\epsilon_0}\ln\frac{r_0}{r}$$

Here r_0 is the arbitrary, *finite* radius at which $V = 0$. We take $r_0 = r_b$, the radius of the inner surface of the outer cylinder. Then the potential at the outer surface of the inner cylinder (at which $r = r_a$) is just the potential V_{ab} of the inner (positive) cylinder a with respect to the outer (negative) cylinder b:

Figure **24.6** A long cylindrical capacitor. The linear charge density λ is assumed to be positive in this figure. The magnitude of charge in a length L of either cylinder is λL.



$$V_{ab} = \frac{\lambda}{2\pi\epsilon_0}\ln\frac{r_b}{r_a}$$

If λ is positive as in Fig. 24.6, then V_{ab} is positive as well: The inner cylinder is at higher potential than the outer.

Continued

The total charge Q in a length L is $Q = \lambda L$, so from Eq. (24.1) the capacitance C of a length L is

$$C = \frac{Q}{V_{ab}} = \frac{\lambda L}{\dfrac{\lambda}{2\pi\epsilon_0}\ln\dfrac{r_b}{r_a}} = \frac{2\pi\epsilon_0 L}{\ln(r_b/r_a)}$$

The capacitance per unit length is

$$\frac{C}{L} = \frac{2\pi\epsilon_0}{\ln(r_b/r_a)}$$

Substituting $\epsilon_0 = 8.85 \times 10^{-12}\ \mathrm{F/m} = 8.85\ \mathrm{pF/m}$, we get

$$\frac{C}{L} = \frac{55.6\ \mathrm{pF/m}}{\ln(r_b/r_a)}$$

EVALUATE The capacitance of coaxial cylinders is determined entirely by their dimensions, just as for parallel-plate and spherical capacitors. Ordinary coaxial cables are made like this but with an insulating material instead of vacuum between the conductors. A typical cable used for connecting a television to a cable TV feed has a capacitance per unit length of 69 pF/m.

KEYCONCEPT The capacitance per unit length of two long, coaxial, conducting cylinders depends on the ratio of the inner radius of the outer conductor to the outer radius of the inner conductor.

Figure 24.7 An assortment of commercially available capacitors.

TEST YOUR UNDERSTANDING OF SECTION 24.1 A capacitor has vacuum in the space between the conductors. If you double the amount of charge on each conductor, what happens to the capacitance? (i) It increases; (ii) it decreases; (iii) it remains the same; (iv) the answer depends on the size or shape of the conductors.

ANSWER

(iii) The capacitance does not depend on the value of the charge Q. Doubling Q causes the potential difference V_{ab} to double, so the capacitance $C = Q/V_{ab}$ remains the same. These statements are true no matter what the geometry of the capacitor.

24.2 CAPACITORS IN SERIES AND PARALLEL

Capacitors are manufactured with certain standard capacitances and working voltages (**Fig. 24.7**). However, these standard values may not be the ones you actually need in a particular application. You can obtain the values you need by combining capacitors; many combinations are possible, but the simplest combinations are a series connection and a parallel connection.

Capacitors in Series

Figure 24.8 A series connection of two capacitors.

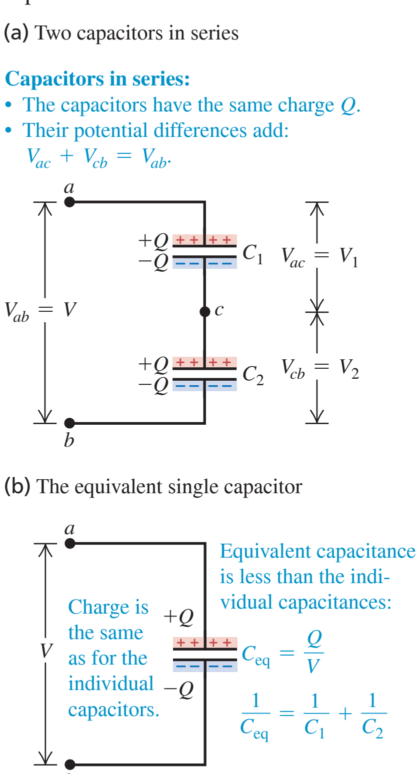


Figure 24.8a is a schematic diagram of a **series connection.** Two capacitors are connected in series (one after the other) by conducting wires between points a and b. Both capacitors are initially uncharged. When a constant positive potential difference V_{ab} is applied between points a and b, the capacitors become charged; the figure shows that the charge on *all* conducting plates has the same magnitude. To see why, note first that the top plate of C_1 acquires a positive charge Q. The electric field of this positive charge pulls negative charge up to the bottom plate of C_1 until all of the field lines that begin on the top plate end on the bottom plate. This requires that the bottom plate have charge $-Q$. These negative charges had to come from the top plate of C_2, which becomes positively charged with charge $+Q$. This positive charge then pulls negative charge $-Q$ from the connection at point b onto the bottom plate of C_2. The total charge on the lower plate of C_1 and the upper plate of C_2 together must always be zero because these plates aren't connected to anything except each other. Thus *in a series connection the magnitude of charge on all plates is the same.*

Referring to Fig. 24.8a, we can write the potential differences between points a and c, c and b, and a and b as

$$V_{ac} = V_1 = \frac{Q}{C_1}, \qquad V_{cb} = V_2 = \frac{Q}{C_2}, \qquad V_{ab} = V = V_1 + V_2 = Q\left(\frac{1}{C_1} + \frac{1}{C_2}\right)$$

and so

$$\frac{V}{Q} = \frac{1}{C_1} + \frac{1}{C_2} \tag{24.3}$$

Following a common convention, we use the symbols V_1, V_2, and V to denote the potential *differences* V_{ac} (across the first capacitor), V_{cb} (across the second capacitor), and V_{ab} (across the entire combination of capacitors), respectively.

The **equivalent capacitance** C_{eq} of the series combination is defined as the capacitance of a *single* capacitor for which the charge Q is the same as for the combination, when the potential difference V is the same. In other words, the combination can be replaced by an *equivalent capacitor* of capacitance C_{eq}. For such a capacitor, shown in Fig. 24.8b,

$$C_{eq} = \frac{Q}{V} \quad \text{or} \quad \frac{1}{C_{eq}} = \frac{V}{Q} \tag{24.4}$$

Combining Eqs. (24.3) and (24.4), we find

$$\frac{1}{C_{eq}} = \frac{1}{C_1} + \frac{1}{C_2}$$

We can extend this analysis to any number of capacitors in series. We find the following result for the *reciprocal* of the equivalent capacitance:

Capacitors in series:

$$\underbrace{\frac{1}{C_{eq}}}_{\text{Equivalent capacitance of series combination}} = \underbrace{\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}}_{\text{Capacitances of individual capacitors}} + \cdots \tag{24.5}$$

The reciprocal of the equivalent capacitance of a series combination equals the sum of the reciprocals of the individual capacitances. In a series connection the equivalent capacitance is always *less than* any individual capacitance.

CAUTION **Capacitors in series** The magnitude of charge is the same on all plates of all the capacitors in a series combination; however, the potential differences of the individual capacitors are *not* the same unless their individual capacitances are the same. The potential differences of the individual capacitors add to give the total potential difference across the series combination: $V_{total} = V_1 + V_2 + V_3 + \cdots$.

APPLICATION **Touch Screens and Capacitance** The touch screen on a mobile phone, a tablet, or (as shown here) a medical device uses the physics of capacitors. Behind the screen are two parallel layers, one behind the other, of thin strips of a transparent conductor such as indium tin oxide. A voltage is maintained between the two layers. The strips in one layer are oriented perpendicular to those in the other layer; the points where two strips overlap act as a grid of capacitors. When you bring your finger (a conductor) up to a point on the screen, your finger and the front conducting layer act like a second capacitor in series at that point. The circuitry attached to the conducting layers detects the location of the capacitance change, and so detects where you touched the screen.

Capacitors in Parallel

The arrangement shown in **Fig. 24.9a** is called a **parallel connection.** Two capacitors are connected in parallel between points a and b. In this case the upper plates of the two capacitors are connected by conducting wires to form an equipotential surface, and the lower plates form another. Hence *in a parallel connection the potential difference for all individual capacitors is the same* and is equal to $V_{ab} = V$. The charges Q_1 and Q_2 are not necessarily equal, however, since charges can reach each capacitor independently from the source (such as a battery) of the voltage V_{ab}. The charges are

$$Q_1 = C_1 V \quad \text{and} \quad Q_2 = C_2 V$$

The *total* charge Q of the combination, and thus the total charge on the equivalent capacitor, is

$$Q = Q_1 + Q_2 = (C_1 + C_2)V$$

so

$$\frac{Q}{V} = C_1 + C_2 \tag{24.6}$$

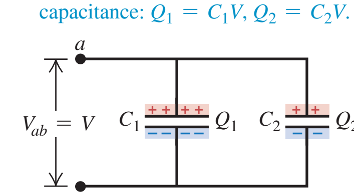
The parallel combination is equivalent to a single capacitor with the same total charge $Q = Q_1 + Q_2$ and potential difference V as the combination (Fig. 24.9b). The equivalent capacitance of the combination, C_{eq}, is the same as the capacitance Q/V of this single equivalent capacitor. So from Eq. (24.6),

$$C_{eq} = C_1 + C_2$$

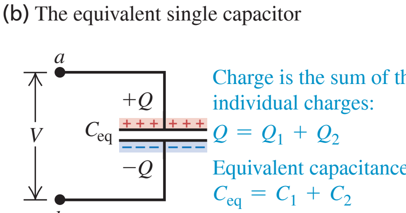
Figure **24.9** A parallel connection of two capacitors.

(a) Two capacitors in parallel

Capacitors in parallel:
- The capacitors have the same potential V.
- The charge on each capacitor depends on its capacitance: $Q_1 = C_1V$, $Q_2 = C_2V$.



(b) The equivalent single capacitor



In the same way we can show that for any number of capacitors in parallel,

Capacitors in parallel:

$$C_{eq} = C_1 + C_2 + C_3 + \cdots \quad (24.7)$$

Equivalent capacitance of parallel combination

Capacitances of individual capacitors

The equivalent capacitance of a parallel combination equals the *sum* of the individual capacitances. In a parallel connection the equivalent capacitance is always *greater than* any individual capacitance.

CAUTION **Capacitors in parallel** The potential differences are the same for all capacitors in a parallel combination; however, the charges on individual capacitors are *not* the same unless their individual capacitances are the same. The charges on the individual capacitors add to give the total charge on the parallel combination: $Q_{total} = Q_1 + Q_2 + Q_3 + \cdots$. [Compare these statements to those in the "Caution" paragraph following Eq. (24.5).]

PROBLEM-SOLVING STRATEGY 24.1 Equivalent Capacitance

IDENTIFY *the relevant concepts:* The concept of equivalent capacitance is useful whenever two or more capacitors are connected.

SET UP *the problem* using the following steps:

1. Make a drawing of the capacitor arrangement.
2. Identify all groups of capacitors that are connected in series or in parallel.
3. Keep in mind that when we say a capacitor "has charge Q," we mean that the plate at higher potential has charge $+Q$ and the other plate has charge $-Q$.

EXECUTE *the solution* as follows:

1. Use Eq. (24.5) to find the equivalent capacitance of capacitors connected in series, as in Fig. 24.8. Such capacitors each have the *same charge* if they were uncharged before they were connected; that charge is the same as that on the equivalent capacitor. The potential difference across the combination is the sum of the potential differences across the individual capacitors.
2. Use Eq. (24.7) to find the equivalent capacitance of capacitors connected in parallel, as in Fig. 24.9. Such capacitors all have the *same potential difference* across them; that potential difference is the same as that across the equivalent capacitor. The total charge on the combination is the sum of the charges on the individual capacitors.
3. After replacing all the series or parallel groups you initially identified, you may find that more such groups reveal themselves. Replace those groups by using the same procedure as above until no more replacements are possible. If you then need to find the charge or potential difference for an individual original capacitor, you may have to retrace your steps.

EVALUATE *your answer:* Check whether your result makes sense. If the capacitors are connected in series, the equivalent capacitance C_{eq} must be *smaller* than any of the individual capacitances. If the capacitors are connected in parallel, C_{eq} must be *greater* than any of the individual capacitances.

EXAMPLE 24.5 Capacitors in series and in parallel

WITH VARIATION PROBLEMS

In Figs. 24.8 and 24.9, let $C_1 = 6.0\ \mu\text{F}$, $C_2 = 3.0\ \mu\text{F}$, and $V_{ab} = 18\ \text{V}$. Find the equivalent capacitance and the charge and potential difference for each capacitor when the capacitors are connected (a) in series (see Fig. 24.8) and (b) in parallel (see Fig. 24.9).

IDENTIFY and SET UP In both parts of this example a target variable is the equivalent capacitance C_{eq}, which is given by Eq. (24.5) for the series combination in part (a) and by Eq. (24.7) for the parallel combination in part (b). In each part we find the charge and potential difference from the definition of capacitance, Eq. (24.1), and the rules outlined in Problem-Solving Strategy 24.1.

EXECUTE (a) From Eq. (24.5) for a series combination,

$$\frac{1}{C_{eq}} = \frac{1}{C_1} + \frac{1}{C_2} = \frac{1}{6.0\ \mu\text{F}} + \frac{1}{3.0\ \mu\text{F}} \qquad C_{eq} = 2.0\ \mu\text{F}$$

The charge Q on each capacitor in series is the same as that on the equivalent capacitor:

$$Q = C_{eq}V = (2.0\ \mu\text{F})(18\ \text{V}) = 36\ \mu\text{C}$$

The potential difference across each capacitor is inversely proportional to its capacitance:

$$V_{ac} = V_1 = \frac{Q}{C_1} = \frac{36\ \mu\text{C}}{6.0\ \mu\text{F}} = 6.0\ \text{V}$$

$$V_{cb} = V_2 = \frac{Q}{C_2} = \frac{36\ \mu\text{C}}{3.0\ \mu\text{F}} = 12.0\ \text{V}$$

(b) From Eq. (24.7) for a parallel combination,

$$C_{eq} = C_1 + C_2 = 6.0\ \mu\text{F} + 3.0\ \mu\text{F} = 9.0\ \mu\text{F}$$

The potential difference across each of the capacitors is the same as that across the equivalent capacitor, 18 V. The charge on each capacitor is directly proportional to its capacitance:

$$Q_1 = C_1V = (6.0\ \mu\text{F})(18\ \text{V}) = 108\ \mu\text{C}$$

$$Q_2 = C_2V = (3.0\ \mu\text{F})(18\ \text{V}) = 54\ \mu\text{C}$$

EVALUATE As expected, the equivalent capacitance C_{eq} for the series combination in part (a) is less than either C_1 or C_2, while that for the parallel combination in part (b) is greater than either C_1 or C_2. For two capacitors in series, as in part (a), the charge is the same on either capacitor and the *larger* potential difference appears across the capacitor with the *smaller* capacitance. Furthermore, the sum of the potential differences across the individual capacitors in series equals the potential difference across the equivalent capacitor: $V_{ac} + V_{cb} = V_{ab} = 18\text{ V}$. By contrast, for two capacitors in parallel, as in part (b), each capacitor has the same potential difference and the *larger* charge appears on the capacitor with the *larger* capacitance. Can you show that the total charge $Q_1 + Q_2$ on the parallel combination is equal to the charge $Q = C_{eq}V$ on the equivalent capacitor?

KEYCONCEPT When capacitors are combined in *series,* the charge is the same on each capacitor but the potential differences across the capacitors may not be the same. The reciprocal of the equivalent capacitance of a series combination equals the sum of the reciprocals of the individual capacitances. When capacitors are combined in *parallel,* the potential difference across each capacitor is the same but the charges on the capacitors may not be the same. The equivalent capacitance of a parallel combination equals the sum of the individual capacitances.

EXAMPLE 24.6 A capacitor network

WITH VARIATION PROBLEMS

Find the equivalent capacitance of the five-capacitor network shown in **Fig. 24.10a**.

IDENTIFY and SET UP These capacitors are neither all in series nor all in parallel. We can, however, identify portions of the arrangement that *are* either in series or parallel. We combine these as described in Problem-Solving Strategy 24.1 to find the net equivalent capacitance, using Eq. (24.5) for series connections and Eq. (24.7) for parallel connections.

EXECUTE The caption of Fig. 24.10 outlines our procedure. We first use Eq. (24.5) to replace the 12 μF and 6 μF series combination by its equivalent capacitance C':

$$\frac{1}{C'} = \frac{1}{12\ \mu\text{F}} + \frac{1}{6\ \mu\text{F}} \qquad C' = 4\ \mu\text{F}$$

This gives us the equivalent combination of Fig. 24.10b. Now we see three capacitors in parallel, and we use Eq. (24.7) to replace them with their equivalent capacitance C'':

$$C'' = 3\ \mu\text{F} + 11\ \mu\text{F} + 4\ \mu\text{F} = 18\ \mu\text{F}$$

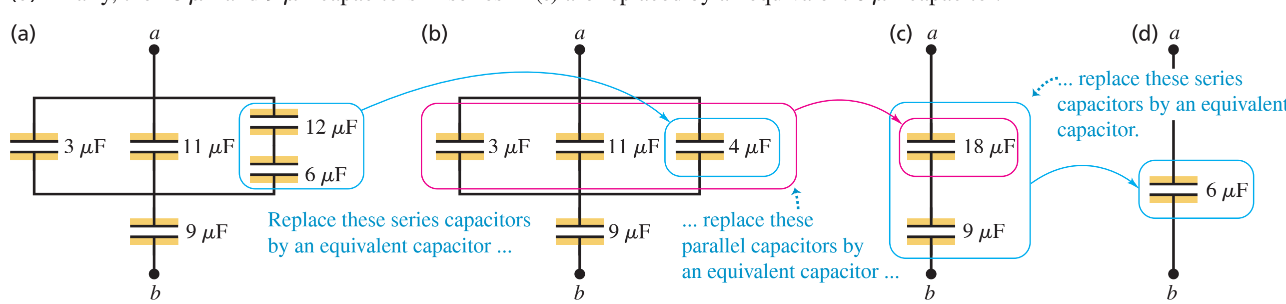
This gives us the equivalent combination of Fig. 24.10c, which has two capacitors in series. We use Eq. (24.5) to replace them with their equivalent capacitance C_{eq}, which is our target variable (Fig. 24.10d):

$$\frac{1}{C_{eq}} = \frac{1}{18\ \mu\text{F}} + \frac{1}{9\ \mu\text{F}} \qquad C_{eq} = 6\ \mu\text{F}$$

EVALUATE If the potential difference across the entire network in Fig. 24.10a is $V_{ab} = 9.0\text{ V}$, the net charge on the network is $Q = C_{eq}V_{ab} = (6\ \mu\text{F})(9.0\text{ V}) = 54\ \mu\text{C}$. Can you find the charge on, and the voltage across, each of the five individual capacitors?

KEYCONCEPT To analyze any network of capacitors, first identify which portions of the network are in series and which are in parallel. Then use the rules for series and parallel capacitors to determine the equivalent capacitance of the network.

Figure 24.10 (a) A capacitor network between points a and b. (b) The 12 μF and 6 μF capacitors in series in (a) are replaced by an equivalent 4 μF capacitor. (c) The 3 μF, 11 μF, and 4 μF capacitors in parallel in (b) are replaced by an equivalent 18 μF capacitor. (d) Finally, the 18 μF and 9 μF capacitors in series in (c) are replaced by an equivalent 6 μF capacitor.



TEST YOUR UNDERSTANDING OF SECTION 24.2 (a) How should you connect a 4 μF capacitor and an 8 μF capacitor so that the 4 μF capacitor has a greater potential difference across it than the 8 μF capacitor? (i) Series; (ii) parallel; (iii) either series or parallel; (iv) neither series nor parallel. (b) How should you connect them so that the 4 μF capacitor has a greater charge than the 8 μF capacitor? (i) Series; (ii) parallel; (iii) either series or parallel; (iv) neither series nor parallel.

ANSWER

(a) **(i)**, (b) **(iv)** In a series connection the two capacitors carry the same charge Q but have different potential differences $V_{ab} = Q/C$; the capacitor with the smaller capacitance C has the greater potential difference. In a parallel connection the two capacitors have the same potential difference V_{ab} but carry different charges $Q = CV_{ab}$; the capacitor with the larger capacitance C has the greater charge. Hence a 4 μF capacitor will have a greater potential difference than an 8 μF capacitor if the two are connected in series. The 4 μF capacitor cannot carry more charge than the 8 μF capacitor no matter how they are connected: In a series connection they will carry the same charge, and in a parallel connection the 8 μF capacitor will carry more charge.

24.3 ENERGY STORAGE IN CAPACITORS AND ELECTRIC-FIELD ENERGY

Many of the most important applications of capacitors depend on their ability to store energy. The electric potential energy stored in a charged capacitor is just equal to the amount of work required to charge it—that is, to separate opposite charges and place them on different conductors. When the capacitor is discharged, this stored energy is recovered as work done by electrical forces.

We can calculate the potential energy U of a charged capacitor by calculating the work W required to charge it. Suppose that when we are done charging the capacitor, the final charge is Q and the final potential difference is V. From Eq. (24.1) these quantities are related by

$$V = \frac{Q}{C}$$

Let q and v be the charge and potential difference, respectively, at an intermediate stage during the charging process; then $v = q/C$. At this stage the work dW required to transfer an additional element of charge dq is

$$dW = v\,dq = \frac{q\,dq}{C}$$

The total work W needed to increase the capacitor charge q from zero to Q is

$$W = \int_0^W dW = \frac{1}{C}\int_0^Q q\,dq = \frac{Q^2}{2C} \qquad \text{(work to charge a capacitor)} \qquad (24.8)$$

This is also the total work done by the electric field on the charge when the capacitor discharges. Then q *decreases* from an initial value Q to zero as the elements of charge dq "fall" through potential differences v that vary from V down to zero.

If we define the potential energy of an *uncharged* capacitor to be zero, then W in Eq. (24.8) is equal to the potential energy U of the charged capacitor. The final stored charge is $Q = CV$, so we can express U (which is equal to W) as

Potential energy stored in a capacitor Magnitude of charge on each plate

$$U = \frac{Q^2}{2C} = \tfrac{1}{2}CV^2 = \tfrac{1}{2}QV \qquad (24.9)$$

Capacitance Potential difference between plates

When Q is in coulombs, C in farads (coulombs per volt), and V in volts (joules per coulomb), U is in joules.

The last form of Eq. (24.9), $U = \frac{1}{2}QV$, shows that the total work W required to charge the capacitor is equal to the total charge Q multiplied by the *average* potential difference $\frac{1}{2}V$ during the charging process.

The expression $U = \frac{1}{2}(Q^2/C)$ in Eq. (24.9) shows that a charged capacitor is the electrical analog of a stretched spring with elastic potential energy $U = \frac{1}{2}kx^2$. The charge Q is analogous to the elongation x, and the *reciprocal* of the capacitance, $1/C$, is analogous to the force constant k. The energy supplied to a capacitor in the charging process is analogous to the work we do on a spring when we stretch it.

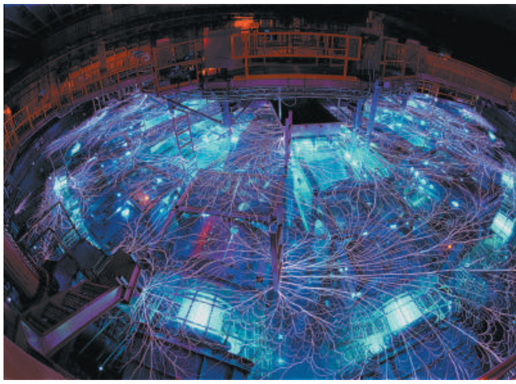
Equations (24.8) and (24.9) tell us that capacitance measures the ability of a capacitor to store both energy and charge. If a capacitor is charged by connecting it to a battery or other source that provides a fixed potential difference V, then increasing the value of C gives a greater charge $Q = CV$ and a greater amount of stored energy $U = \frac{1}{2}CV^2$. If instead the goal is to transfer a given quantity of charge Q from one conductor to another, Eq. (24.8) shows that the work W required is inversely proportional to C; the greater the capacitance, the easier it is to give a capacitor a fixed amount of charge.

Applications of Capacitors: Energy Storage

Most practical applications of capacitors take advantage of their ability to store and release energy. In an electronic flash unit used in photography, the energy stored in a capacitor (see Fig. 24.4) is released when the button is pressed to take a photograph. This provides a conducting path from one capacitor plate to the other through the flash tube. Once this path is established, the stored energy is rapidly converted into a brief but intense flash of light. An extreme example of the same principle is the Z machine at Sandia National Laboratories in New Mexico, which is used in experiments in controlled nuclear fusion (**Fig. 24.11**). A bank of charged capacitors releases more than a million joules of energy in just a few billionths of a second. For that brief space of time, the power output of the Z machine is 2.9×10^{14} W, or about 80 times the power output of all the electric power plants on earth combined!

In other applications, the energy is released more slowly. Springs in the suspension of an automobile help smooth out the ride by absorbing the energy from sudden jolts and releasing that energy gradually; in an analogous way, a capacitor in an electronic circuit can smooth out unwanted variations in voltage due to power surges. We'll discuss these circuits in detail in Chapter 26.

Figure 24.11 The Z machine uses a large number of capacitors in parallel to give a tremendous equivalent capacitance C (see Section 24.2). Hence a large amount of energy $U = \frac{1}{2}CV^2$ can be stored with even a modest potential difference V. The arcs shown here are produced when the capacitors discharge their energy into a target, which is no larger than a spool of thread. This heats the target to a temperature higher than 2×10^9 K.



Electric-Field Energy

We can charge a capacitor by moving electrons directly from one plate to another. This requires doing work against the electric field between the plates. Thus we can think of the energy as being stored *in the field* in the region between the plates. To see this, let's find the energy *per unit volume* in the space between the plates of a parallel-plate capacitor with plate area A and separation d. We call this the **energy density,** denoted by u. From Eq. (24.9) the total stored potential energy is $\frac{1}{2}CV^2$ and the volume between the plates is Ad; hence

$$u = \text{Energy density} = \frac{\frac{1}{2}CV^2}{Ad} \tag{24.10}$$

From Eq. (24.2) the capacitance C is given by $C = \epsilon_0 A/d$. The potential difference V is related to the electric-field magnitude E by $V = Ed$. If we use these expressions in Eq. (24.10), the geometric factors A and d cancel, and we find

Electric energy density in a vacuum

$$u = \frac{1}{2}\epsilon_0 E^2 \tag{24.11}$$

Magnitude of electric field

Electric constant

Although we have derived this relationship for a parallel-plate capacitor only, it turns out to be valid for any capacitor in vacuum and indeed *for any electric-field configuration in vacuum*. This result has an interesting implication. We think of vacuum as space with no matter in it, but vacuum can nevertheless have electric fields and therefore energy. Thus "empty" space need not be truly empty after all. We'll use this idea and Eq. (24.11) in Chapter 32 in connection with the energy transported by electromagnetic waves.

CAUTION **Electric-field energy is electric potential energy** It's a common misconception that electric-field energy is a new kind of energy, different from the electric potential energy described before. This is *not* the case; it is simply a different way of interpreting electric potential energy. We can regard the energy of a given system of charges as being a shared property of all the charges, or we can think of the energy as being a property of the electric field that the charges create. Either interpretation leads to the same value of the potential energy.

EXAMPLE 24.7 Transferring charge and energy between capacitors

WITH VARIATION PROBLEMS

We connect a capacitor $C_1 = 8.0\ \mu\text{F}$ to a power supply, charge it to a potential difference $V_0 = 120\ \text{V}$, and disconnect the power supply (**Fig. 24.12**). Switch S is open. (a) What is the charge Q_0 on C_1? (b) What is the energy stored in C_1? (c) Capacitor $C_2 = 4.0\ \mu\text{F}$ is initially uncharged. We close switch S. After charge no longer flows, what is the potential difference across each capacitor, and what is the charge on each capacitor? (d) What is the final energy of the system?

IDENTIFY and SET UP In parts (a) and (b) we find the charge Q_0 and stored energy U_{initial} for the single charged capacitor C_1 from Eqs. (24.1) and (24.9), respectively. After we close switch S, one wire connects the upper plates of the two capacitors and another wire connects the lower plates; the capacitors are now connected in parallel. In part (c) we use the character of the parallel connection to determine how Q_0 is shared between the two capacitors. In part (d) we again use Eq. (24.9) to find the energy stored in capacitors C_1 and C_2; the energy of the system is the sum of these values.

EXECUTE (a) The initial charge Q_0 on C_1 is

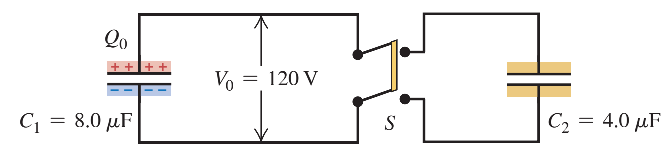
$$Q_0 = C_1 V_0 = (8.0\ \mu\text{F})(120\ \text{V}) = 960\ \mu\text{C}$$

(b) The energy initially stored in C_1 is

$$U_{\text{initial}} = \tfrac{1}{2}Q_0V_0 = \tfrac{1}{2}(960 \times 10^{-6}\ \text{C})(120\ \text{V}) = 0.058\ \text{J}$$

(c) When we close the switch, the positive charge Q_0 is distributed over the upper plates of both capacitors and the negative charge $-Q_0$ is distributed over the lower plates. Let Q_1 and Q_2 be the magnitudes of the final charges on the capacitors. Conservation of charge requires that $Q_1 + Q_2 = Q_0$. The potential difference V between the plates is the same for both capacitors because they are connected in parallel, so the charges are $Q_1 = C_1V$ and $Q_2 = C_2V$. We now have three independent equations relating the three unknowns Q_1, Q_2, and V. Solving these, we find

Figure 24.12 When the switch S is closed, the charged capacitor C_1 is connected to an uncharged capacitor C_2. The center part of the switch is an insulating handle; charge can flow only between the two upper terminals and between the two lower terminals.



$$V = \frac{Q_0}{C_1 + C_2} = \frac{960\ \mu\text{C}}{8.0\ \mu\text{F} + 4.0\ \mu\text{F}} = 80\ \text{V}$$

$$Q_1 = 640\ \mu\text{C} \qquad Q_2 = 320\ \mu\text{C}$$

(d) The final energy of the system is

$$U_{\text{final}} = \tfrac{1}{2}Q_1V + \tfrac{1}{2}Q_2V = \tfrac{1}{2}Q_0V$$
$$= \tfrac{1}{2}(960 \times 10^{-6}\ \text{C})(80\ \text{V}) = 0.038\ \text{J}$$

EVALUATE The final energy is less than the initial energy; the difference was converted to energy of some other form. The conductors become a little warmer because of their resistance, and some energy is radiated as electromagnetic waves. We'll study the behavior of capacitors in more detail in Chapters 26 and 31.

KEYCONCEPT There are three useful ways to write the amount of electric potential energy stored in a charged capacitor: in terms of the charge Q on the plates and the potential difference V between the plates, in terms of Q and the capacitance C, and in terms of V and C.

EXAMPLE 24.8 Electric-field energy

WITH VARIATION PROBLEMS

(a) What is the magnitude of the electric field required to store 1.00 J of electric potential energy in a volume of 1.00 m^3 in vacuum? (b) If the field magnitude is 10 times larger than that, how much energy is stored per cubic meter?

IDENTIFY and SET UP We use the relationship between the electric-field magnitude E and the energy density u. In part (a) we use the given information to find u; then we use Eq. (24.11) to find the corresponding value of E. In part (b), Eq. (24.11) tells us how u varies with E.

EXECUTE (a) The desired energy density is $u = 1.00\ \text{J/m}^3$. Then from Eq. (24.11),

$$E = \sqrt{\frac{2u}{\epsilon_0}} = \sqrt{\frac{2(1.00\ \text{J/m}^3)}{8.85 \times 10^{-12}\ \text{C}^2/\text{N}\cdot\text{m}^2}}$$
$$= 4.75 \times 10^5\ \text{N/C} = 4.75 \times 10^5\ \text{V/m}$$

(b) Equation (24.11) shows that u is proportional to E^2. If E increases by a factor of 10, u increases by a factor of $10^2 = 100$, so the energy density becomes $u = 100\ \text{J/m}^3$.

EVALUATE Dry air can sustain an electric field of about 3×10^6 V/m without experiencing *dielectric breakdown,* which we'll discuss in Section 24.4. There we'll see that field magnitudes in practical insulators can be even larger than this.

KEYCONCEPT Energy is stored in any region of space where an electric field is present. The electric energy density (energy per unit volume) at a point is proportional to the square of the electric-field magnitude at that point.

EXAMPLE 24.9 Two ways to calculate energy stored in a capacitor

WITH VARIATION PROBLEMS

The spherical capacitor described in Example 24.3 (Section 24.1) has charges $+Q$ and $-Q$ on its inner and outer conductors. Find the electric potential energy stored in the capacitor (a) by using the capacitance C found in Example 24.3 and (b) by integrating the electric-field energy density u.

IDENTIFY and SET UP We can determine the energy U stored in a capacitor in two ways: in terms of the work done to put the charges on the two conductors, and in terms of the energy in the electric field between the conductors. The descriptions are equivalent, so they must give us the same result. In Example 24.3 we found the capacitance C and the field magnitude E in the space between the conductors. (The electric field is zero inside the inner sphere and is also zero outside the inner surface of the outer sphere, because a Gaussian surface with radius $r < r_a$ or $r > r_b$ encloses zero net charge. Hence the energy density is nonzero only in the space between the spheres, $r_a < r < r_b$.) In part (a) we use Eq. (24.9) to find U. In part (b) we use Eq. (24.11) to find u, which we integrate over the volume between the spheres to find U.

EXECUTE (a) From Example 24.3, the spherical capacitor has capacitance

$$C = 4\pi\epsilon_0 \frac{r_a r_b}{r_b - r_a}$$

where r_a and r_b are the radii of the inner and outer conducting spheres, respectively. From Eq. (24.9) the energy stored in this capacitor is

$$U = \frac{Q^2}{2C} = \frac{Q^2}{8\pi\epsilon_0}\frac{r_b - r_a}{r_a r_b}$$

(b) The electric field in the region $r_a < r < r_b$ between the two conducting spheres has magnitude $E = Q/4\pi\epsilon_0 r^2$. The energy density in this region is

$$u = \tfrac{1}{2}\epsilon_0 E^2 = \tfrac{1}{2}\epsilon_0\left(\frac{Q}{4\pi\epsilon_0 r^2}\right)^2 = \frac{Q^2}{32\pi^2\epsilon_0 r^4}$$

The energy density is *not* uniform; it decreases rapidly with increasing distance from the center of the capacitor. To find the total electric-field energy, we integrate u (the energy per unit volume) over the region $r_a < r < r_b$. We divide this region into spherical shells of radius r, surface area $4\pi r^2$, thickness dr, and volume $dV = 4\pi r^2\,dr$. Then

$$U = \int u\,dV = \int_{r_a}^{r_b}\left(\frac{Q^2}{32\pi^2\epsilon_0 r^4}\right)4\pi r^2\,dr$$
$$= \frac{Q^2}{8\pi\epsilon_0}\int_{r_a}^{r_b}\frac{dr}{r^2} = \frac{Q^2}{8\pi\epsilon_0}\left(-\frac{1}{r_b} + \frac{1}{r_a}\right) = \frac{Q^2}{8\pi\epsilon_0}\frac{r_b - r_a}{r_a r_b}$$

EVALUATE Electric potential energy can be associated with either the *charges,* as in part (a), or the *field,* as in part (b); the calculated amount of stored energy is the same in either case.

KEYCONCEPT An alternative way to calculate the electric potential energy stored in a capacitor is to first find the electric field at each point in the space between the conducting plates. From this, find the electric energy density at each point. Finally, integrate the electric energy density over the entire volume where the electric field is present.

TEST YOUR UNDERSTANDING OF SECTION 24.3 You want to connect a 4 μF capacitor and an 8 μF capacitor. With which type of connection will the 4 μF capacitor have a greater amount of *stored energy* than the 8 μF capacitor? (i) Series; (ii) parallel; (iii) either series or parallel; (iv) neither series nor parallel.

ANSWER

(i) Capacitors connected in series carry the same charge Q. To compare the amount of energy stored, we use the expression $U = Q^2/2C$ from Eq. (24.9); it shows that the capacitor with the *smaller* capacitance ($C = 4\ \mu$F) has more stored energy in a series combination. By contrast, capacitors in parallel have the same potential difference V, so to compare them we use $U = \frac{1}{2}CV^2$ from Eq. (24.9). It shows that in a parallel combination, the capacitor with the *larger* capacitance ($C = 8\ \mu$F) has more stored energy. (If we had instead used $U = \frac{1}{2}CV^2$ to analyze the series combination, we would have to account for the different potential differences across the two capacitors. Likewise, using $U = Q^2/2C$ to study the parallel combination would require us to account for the different charges on the capacitors.)

24.4 DIELECTRICS

Most capacitors have a nonconducting material, or **dielectric,** between their conducting plates. A common type of capacitor uses long strips of metal foil for the plates, separated by strips of plastic sheet such as Mylar. A sandwich of these materials is rolled up, forming a unit that can provide a capacitance of several microfarads in a compact package (**Fig. 24.13**).

Placing a solid dielectric between the plates of a capacitor serves three functions. First, it solves the mechanical problem of maintaining two large metal sheets at a very small separation without actual contact.

Second, using a dielectric increases the maximum possible potential difference between the capacitor plates. Any insulating material, when subjected to a sufficiently large electric field, experiences a partial ionization that permits conduction through it (Section 23.3).

Figure 24.13 A common type of capacitor uses dielectric sheets to separate the conductors.

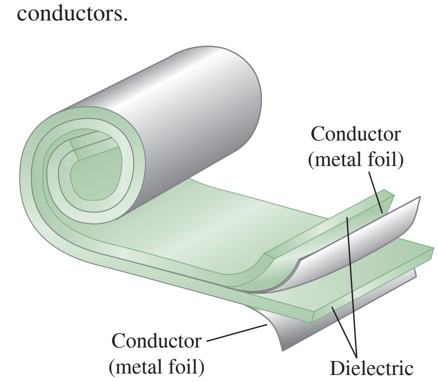
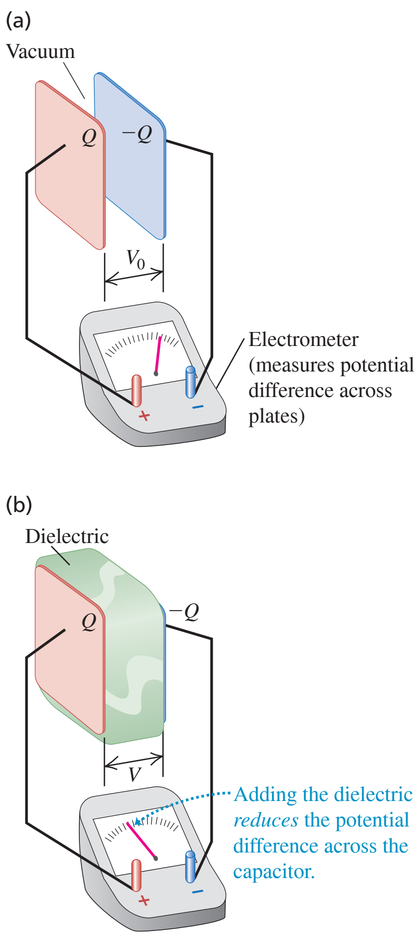


Figure 24.14 Effect of a dielectric between the plates of a parallel-plate capacitor. (a) With a given charge, the potential difference is V_0. (b) With the same charge but with a dielectric between the plates, the potential difference V is smaller than V_0.



This phenomenon is called **dielectric breakdown.** Many dielectric materials can tolerate stronger electric fields without breakdown than can air. Thus using a dielectric allows a capacitor to sustain a higher potential difference V and so store greater amounts of charge and energy.

Third, the capacitance of a capacitor of given dimensions is *greater* when there is a dielectric material between the plates than when there is vacuum. We can demonstrate this effect with the aid of a sensitive *electrometer,* a device that measures the potential difference between two conductors without letting any appreciable charge flow from one to the other. **Figure 24.14a** shows an electrometer connected across a charged capacitor, with magnitude of charge Q on each plate and potential difference V_0. When we insert an uncharged sheet of dielectric, such as glass, paraffin, or polystyrene, between the plates, experiment shows that the potential difference *decreases* to a smaller value V (Fig. 24.14b). When we remove the dielectric, the potential difference returns to its original value V_0, showing that the original charges on the plates have not changed.

The original capacitance C_0 is given by $C_0 = Q/V_0$, and the capacitance C with the dielectric present is $C = Q/V$. The charge Q is the same in both cases, and V is less than V_0, so we conclude that the capacitance C with the dielectric present is *greater* than C_0. When the space between plates is completely filled by the dielectric, the ratio of C to C_0 (equal to the ratio of V_0 to V) is called the **dielectric constant** of the material, K:

$$K = \frac{C}{C_0} \qquad \text{(definition of dielectric constant)} \tag{24.12}$$

When the charge is constant, $Q = C_0V_0 = CV$ and $C/C_0 = V_0/V$. In this case,

$$V = \frac{V_0}{K} \qquad \text{(when } Q \text{ is constant)} \tag{24.13}$$

With the dielectric present, the potential difference for a given charge Q is *reduced* by a factor K.

The dielectric constant K is a pure number. Because C is always greater than C_0, K is always greater than unity. **Table 24.1** gives some representative values of K. For vacuum, $K = 1$ by definition. For air at ordinary temperatures and pressures, K is about 1.0006; this is so nearly equal to 1 that for most purposes an air capacitor is equivalent to one in vacuum. Note that while water has a very large value of K, it is usually not a very practical dielectric for use in capacitors. The reason is that while pure water is a very poor conductor, it is also an excellent ionic solvent. Any ions that are dissolved in the water will cause charge to flow between the capacitor plates, so the capacitor discharges.

CAUTION **Dielectric constant vs. electric constant** Don't confuse the *dielectric* constant K with the *electric* constant ϵ_0. The value of K is a pure number with no units and is different for different materials (see Table 24.1). By contrast, ϵ_0 is a universal constant with units $C^2/N \cdot m^2$ or F/m.

TABLE 24.1 Values of Dielectric Constant K at 20°C

Material	K	Material	K
Vacuum	1	Polyvinyl chloride	3.18
Air (1 atm)	1.00059	Plexiglas®	3.40
Air (100 atm)	1.0548	Glass	5–10
Teflon®	2.1	Neoprene	6.70
Polyethylene	2.25	Germanium	16
Benzene	2.28	Glycerin	42.5
Mica	3–6	Water	80.4
Mylar®	3.1	Strontium titanate	310

No real dielectric is a perfect insulator. Hence there is always some *leakage current* between the charged plates of a capacitor with a dielectric. We tacitly ignored this effect in Section 24.2 when we derived expressions for the equivalent capacitances of capacitors in series, Eq. (24.5), and in parallel, Eq. (24.7). But if a leakage current flows for a long enough time to substantially change the charges from the values we used to derive Eqs. (24.5) and (24.7), those equations may no longer be accurate.

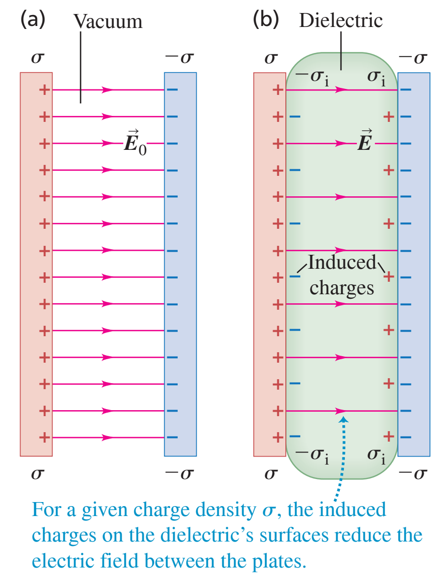
Induced Charge and Polarization

When a dielectric material is inserted between the plates while the charge is kept constant, the potential difference between the plates decreases by a factor K. Therefore the electric field between the plates must decrease by the same factor. If E_0 is the vacuum value and E is the value with the dielectric, then

$$E = \frac{E_0}{K} \qquad \text{(when } Q \text{ is constant)} \tag{24.14}$$

Since the electric-field magnitude is smaller when the dielectric is present, the surface charge density (which causes the field) must be smaller as well. The surface charge on the conducting plates does not change, but an *induced* charge of the opposite sign appears on each surface of the dielectric (**Fig. 24.15**). The dielectric was originally electrically neutral and is still neutral; the induced surface charges arise as a result of *redistribution* of positive and negative charges within the dielectric material, a phenomenon called **polarization.** We first encountered polarization in Section 21.2, and we suggest that you reread the discussion of Fig. 21.8. We'll assume that the induced surface charge is *directly proportional* to the electric-field magnitude E in the material; this is indeed the case for many common dielectrics. (This direct proportionality is analogous to Hooke's law for a spring.) In that case, K is a constant for any particular material. When the electric field is very strong or if the dielectric is made of certain crystalline materials, the relationship between induced charge and the electric field can be more complex; we won't consider such cases here.

Figure **24.15** Electric field lines with (a) vacuum between the plates and (b) dielectric between the plates.



We can derive a relationship between this induced surface charge and the charge on the plates. Let σ_i be the magnitude of the charge per unit area induced on the surfaces of the dielectric (the induced surface charge density). The magnitude of the surface charge density on the capacitor plates is σ, as usual. Then the *net* surface charge on each side of the capacitor has magnitude $(\sigma - \sigma_i)$; see Fig. 24.15b. As we found in Example 21.12 (Section 21.5) and in Example 22.8 (Section 22.4), the field between the plates is related to the net surface charge density by $E = \sigma_{net}/\epsilon_0$. Without and with the dielectric, respectively,

$$E_0 = \frac{\sigma}{\epsilon_0} \qquad E = \frac{\sigma - \sigma_i}{\epsilon_0} \tag{24.15}$$

Using these expressions in Eq. (24.14) and rearranging the result, we find

$$\sigma_i = \sigma\left(1 - \frac{1}{K}\right) \qquad \text{(induced surface charge density)} \tag{24.16}$$

This equation shows that when K is very large, σ_i is nearly as large as σ. In this case, σ_i nearly cancels σ, and the field and potential difference are much smaller than their values in vacuum.

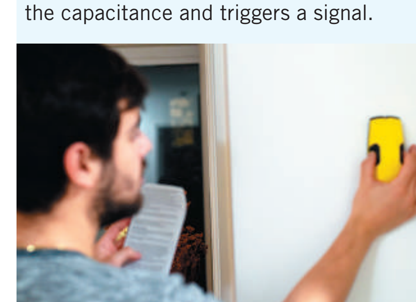
The product $K\epsilon_0$ is called the **permittivity** of the dielectric, denoted by ϵ:

$$\epsilon = K\epsilon_0 \qquad \text{(definition of permittivity)} \tag{24.17}$$

In terms of ϵ we can express the electric field within the dielectric as

$$E = \frac{\sigma}{\epsilon} \tag{24.18}$$

APPLICATION Capacitors in the Toolbox Several practical devices rely on the way a capacitor responds to a change in dielectric constant. One example is an electric stud finder, used to locate metal segments hidden behind a wall's surface. It consists of a metal plate with associated circuitry. The plate acts as one half of a capacitor; the wall acts as the other half. If the stud finder moves over a metal segment, the effective dielectric constant for the capacitor changes, which changes the capacitance and triggers a signal.



Then

Capacitance of a parallel-plate capacitor, dielectric between plates

$$C = \overset{\text{Dielectric constant}}{K}\underset{\text{Capacitance without dielectric}}{C_0} = K\underset{\text{Electric constant}}{\epsilon_0}\frac{\overset{\text{Area of each plate}}{A}}{\underset{\text{Distance between plates}}{d}} = \overset{\text{Permittivity} = K\epsilon_0}{\epsilon}\frac{A}{d} \qquad (24.19)$$

We can repeat the derivation of Eq. (24.11) for the energy density u in an electric field for the case in which a dielectric is present. The result is

Electric energy density in a dielectric

$$u = \tfrac{1}{2}\overset{\text{Dielectric constant}}{K}\underset{\text{Electric constant}}{\epsilon_0}\underset{\text{Magnitude of electric field}}{E^2} = \tfrac{1}{2}\overset{\text{Permittivity} = K\epsilon_0}{\epsilon}E^2 \qquad (24.20)$$

In empty space, where $K = 1$, $\epsilon = \epsilon_0$ and Eqs. (24.19) and (24.20) reduce to Eqs. (24.2) and (24.11), respectively, for a parallel-plate capacitor in vacuum. For this reason, ϵ_0 is sometimes called the "permittivity of free space" or the "permittivity of vacuum." Because K is a pure number, ϵ and ϵ_0 have the same units, $C^2/N \cdot m^2$ or F/m.

Equation (24.19) shows that extremely high capacitances can be obtained with plates that have a large surface area A and are separated by a small distance d by a dielectric with a large value of K. In an *electrolytic double-layer capacitor,* tiny carbon granules adhere to each plate: A is the combined surface area of the granules, which can be tremendous. The plates with granules attached are separated by a very thin dielectric sheet. Such a capacitor can have a capacitance of 5000 F yet fit in the palm of your hand (compare Example 24.1 in Section 24.1).

CAUTION Capacitance is independent of charge or voltage Equation (24.19) shows that the capacitance C of a capacitor depends on only the capacitor geometry (the size, shape, and position of its plates) and whether or not a dielectric is present. The capacitance does *not* depend on the amount of charge placed on the plates or on the voltage between the plates!

PROBLEM-SOLVING STRATEGY 24.2 Dielectrics

IDENTIFY *the relevant concepts:* The relationships in this section are useful whenever there is an electric field in a dielectric, such as a dielectric between charged capacitor plates. Typically you must relate the potential difference V_{ab} between the plates, the electric-field magnitude E in the capacitor, the charge density σ on the capacitor plates, and the induced charge density σ_i on the surfaces of the capacitor.

SET UP *the problem* using the following steps:

1. Make a drawing of the situation.
2. Identify the target variables, and choose which equations from this section will help you solve for those variables.

EXECUTE *the solution* as follows:

1. In problems such as the next example, it is easy to get lost in a blizzard of formulas. Ask yourself at each step what kind of quantity each symbol represents. For example, distinguish clearly between charges and charge densities, and between electric fields and electric potential differences.
2. Check for consistency of units. Distances must be in meters. A microfarad is 10^{-6} farad, and so on. Don't confuse the numerical value of ϵ_0 with the value of $1/4\pi\epsilon_0$. Electric-field magnitude can be expressed in both N/C and V/m. The units of ϵ_0 are $C^2/N \cdot m^2$ or F/m.

EVALUATE *your answer:* With a dielectric present, (a) the capacitance is greater than without a dielectric; (b) for a given charge on the capacitor, the electric field and potential difference are less than without a dielectric; and (c) the magnitude of the induced surface charge density σ_i on the dielectric is less than that of the charge density σ on the capacitor plates.

EXAMPLE 24.10 A capacitor with and without a dielectric

WITH VARIATION PROBLEMS

Suppose the parallel plates in Fig. 24.15 each have an area of $2000\ \text{cm}^2$ ($2.00 \times 10^{-1}\ \text{m}^2$) and are 1.00 cm ($1.00 \times 10^{-2}$ m) apart. We connect the capacitor to a power supply, charge it to a potential difference $V_0 = 3.00$ kV, and disconnect the power supply. We then insert a sheet of insulating plastic material between the plates, completely filling the space between them. We find that the potential difference decreases to 1.00 kV while the charge on each capacitor plate remains constant. Find (a) the original capacitance C_0; (b) the magnitude of charge Q on each plate; (c) the capacitance C after the dielectric is inserted; (d) the dielectric constant K of the dielectric; (e) the permittivity ϵ of the dielectric; (f) the magnitude of the induced charge Q_i on each face of the dielectric; (g) the original electric field E_0 between the plates; and (h) the electric field E after the dielectric is inserted.

IDENTIFY and SET UP This problem uses most of the relationships we have discussed for capacitors and dielectrics. (Energy relationships are treated in Example 24.11.) Most of the target variables can be obtained in several ways. The methods used below are a sample; we encourage you to think of others and compare your results.

EXECUTE (a) With vacuum between the plates, we use Eq. (24.19) with $K = 1$:

$$C_0 = \epsilon_0\frac{A}{d} = (8.85 \times 10^{-12}\ \text{F/m})\frac{2.00 \times 10^{-1}\ \text{m}^2}{1.00 \times 10^{-2}\ \text{m}}$$
$$= 1.77 \times 10^{-10}\ \text{F} = 177\ \text{pF}$$

(b) From the definition of capacitance, Eq. (24.1),

$$Q = C_0V_0 = (1.77 \times 10^{-10}\ \text{F})(3.00 \times 10^3\ \text{V})$$
$$= 5.31 \times 10^{-7}\ \text{C} = 0.531\ \mu\text{C}$$

(c) When the dielectric is inserted, Q is unchanged but the potential difference decreases to $V = 1.00$ kV. Hence from Eq. (24.1), the new capacitance is

$$C = \frac{Q}{V} = \frac{5.31 \times 10^{-7}\ \text{C}}{1.00 \times 10^3\ \text{V}} = 5.31 \times 10^{-10}\ \text{F} = 531\ \text{pF}$$

(d) From Eq. (24.12), the dielectric constant is

$$K = \frac{C}{C_0} = \frac{5.31 \times 10^{-10}\ \text{F}}{1.77 \times 10^{-10}\ \text{F}} = \frac{531\ \text{pF}}{177\ \text{pF}} = 3.00$$

Alternatively, from Eq. (24.13),

$$K = \frac{V_0}{V} = \frac{3000\ \text{V}}{1000\ \text{V}} = 3.00$$

(e) With K from part (d) in Eq. (24.17), the permittivity is

$$\epsilon = K\epsilon_0 = (3.00)(8.85 \times 10^{-12}\ \text{C}^2/\text{N}\cdot\text{m}^2) = 2.66 \times 10^{-11}\ \text{C}^2/\text{N}\cdot\text{m}^2$$

(f) Multiplying both sides of Eq. (24.16) by the plate area A gives the induced charge $Q_\text{i} = \sigma_\text{i}A$ in terms of the charge $Q = \sigma A$ on each plate:

$$Q_\text{i} = Q\left(1 - \frac{1}{K}\right) = (5.31 \times 10^{-7}\ \text{C})\left(1 - \frac{1}{3.00}\right) = 3.54 \times 10^{-7}\text{C}$$

(g) Since the electric field between the plates is uniform, its magnitude is the potential difference divided by the plate separation:

$$E_0 = \frac{V_0}{d} = \frac{3000\ \text{V}}{1.00 \times 10^{-2}\ \text{m}} = 3.00 \times 10^5\ \text{V/m}$$

(h) After the dielectric is inserted,

$$E = \frac{V}{d} = \frac{1000\ \text{V}}{1.00 \times 10^{-2}\ \text{m}} = 1.00 \times 10^5\ \text{V/m}$$

or, from Eq. (24.18),

$$E = \frac{\sigma}{\epsilon} = \frac{Q}{\epsilon A} = \frac{5.31 \times 10^{-7}\ \text{C}}{(2.66 \times 10^{-11}\ \text{C}^2/\text{N}\cdot\text{m}^2)(2.00 \times 10^{-1}\ \text{m}^2)}$$
$$= 1.00 \times 10^5\ \text{V/m}$$

or, from Eq. (24.15),

$$E = \frac{\sigma - \sigma_\text{i}}{\epsilon_0} = \frac{Q - Q_\text{i}}{\epsilon_0 A}$$
$$= \frac{(5.31 - 3.54) \times 10^{-7}\ \text{C}}{(8.85 \times 10^{-12}\ \text{C}^2/\text{N}\cdot\text{m}^2)(2.00 \times 10^{-1}\ \text{m}^2)}$$
$$= 1.00 \times 10^5\ \text{V/m}$$

or, from Eq. (24.14),

$$E = \frac{E_0}{K} = \frac{3.00 \times 10^5\ \text{V/m}}{3.00} = 1.00 \times 10^5\ \text{V/m}$$

EVALUATE Inserting the dielectric increased the capacitance by a factor of $K = 3.00$ and reduced the electric field between the plates by a factor of $1/K = 1/3.00$. It did so by developing induced charges on the faces of the dielectric of magnitude $Q(1 - 1/K) = Q(1 - 1/3.00) = 0.667Q$.

KEYCONCEPT Adding a dielectric (with dielectric constant K) that fills the space between the plates of a capacitor increases the capacitance by a factor of K. For a given amount of charge, adding the dielectric also reduces both the potential difference between the plates and the electric field at each point by a factor of $1/K$.

EXAMPLE 24.11 Energy storage with and without a dielectric

WITH VARIATION PROBLEMS

Find the energy stored in the electric field of the capacitor in Example 24.10 and the energy density, both before and after the dielectric sheet is inserted.

IDENTIFY and SET UP We consider the ideas of energy stored in a capacitor and of electric-field energy density. We use Eq. (24.9) to find the stored energy and Eq. (24.20) to find the energy density.

EXECUTE From Eq. (24.9), the stored energies U_0 and U without and with the dielectric in place are

$$U_0 = \tfrac{1}{2}C_0V_0^2 = \tfrac{1}{2}(1.77 \times 10^{-10}\ \text{F})(3000\ \text{V})^2 = 7.97 \times 10^{-4}\ \text{J}$$
$$U = \tfrac{1}{2}CV^2 = \tfrac{1}{2}(5.31 \times 10^{-10}\ \text{F})(1000\ \text{V})^2 = 2.66 \times 10^{-4}\ \text{J}$$

The final energy is one-third of the original energy.

Equation (24.20) gives the energy densities without and with the dielectric:

$$u_0 = \tfrac{1}{2}\epsilon_0E_0^2 = \tfrac{1}{2}(8.85 \times 10^{-12}\ \text{C}^2/\text{N}\cdot\text{m}^2)(3.00 \times 10^5\ \text{N/C})^2$$
$$= 0.398\ \text{J/m}^3$$

$$u = \tfrac{1}{2}\epsilon E^2 = \tfrac{1}{2}(2.66 \times 10^{-11}\ \text{C}^2/\text{N}\cdot\text{m}^2)(1.00 \times 10^5\ \text{N/C})^2$$
$$= 0.133\ \text{J/m}^3$$

The energy density with the dielectric is one-third of the original energy density.

Continued

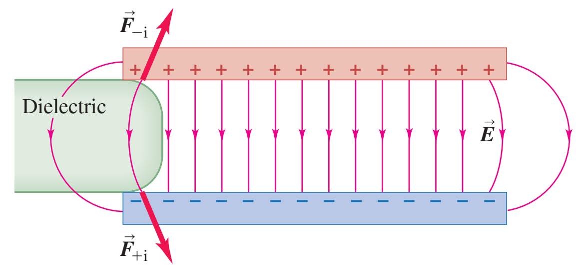
EVALUATE We can check our answer for u_0 by noting that the volume between the plates is $V_{\text{between}} = (0.200\ \text{m}^2)(0.0100\ \text{m}) = 0.00200\ \text{m}^3$. Since the electric field between the plates is uniform, u_0 is uniform as well and the energy density is just the stored energy divided by the volume:

$$u_0 = \frac{U_0}{V_{\text{between}}} = \frac{7.97 \times 10^{-4}\ \text{J}}{0.00200\ \text{m}^3} = 0.398\ \text{J/m}^3$$

This agrees with our earlier answer. You can use the same approach to check our result for u.

In general, when a dielectric is inserted into a capacitor while the charge on each plate remains the same, the permittivity ϵ increases by a factor of K (the dielectric constant), and the electric field E and the energy density $u = \frac{1}{2}\epsilon E^2$ decrease by a factor of $1/K$. Where does the energy go? The answer lies in the fringing field at the edges of a real parallel-plate capacitor. As **Fig. 24.16** shows, that field tends to pull the dielectric into the space between the plates, doing work on it as it does so. We could attach a spring to the left end of the dielectric in Fig. 24.16 and use this force to stretch the spring. Because work is done by the field, the field energy density decreases.

Figure **24.16** The fringing field at the edges of the capacitor exerts forces $\vec{F}_{-\text{i}}$ and $\vec{F}_{+\text{i}}$ on the negative and positive induced surface charges of a dielectric, pulling the dielectric into the capacitor.



KEYCONCEPT Adding a dielectric (with dielectric constant K) that fills the space between the plates of a capacitor reduces the electric field, the electric energy density, and the total stored energy, all by a factor of $1/K$.

Dielectric Breakdown

We mentioned earlier that when a dielectric is subjected to a sufficiently strong electric field, *dielectric breakdown* takes place and the dielectric becomes a conductor. This occurs when the electric field is so strong that electrons are ripped loose from their molecules and crash into other molecules, liberating even more electrons. This avalanche of moving charge forms a spark or arc discharge. Lightning is a dramatic example of dielectric breakdown in air.

Because of dielectric breakdown, capacitors always have maximum voltage ratings. When a capacitor is subjected to excessive voltage, an arc may form through a layer of dielectric, burning or melting a hole in it. This arc creates a conducting path (a short circuit) between the conductors. If a conducting path remains after the arc is extinguished, the device is rendered permanently useless as a capacitor.

The maximum electric-field magnitude that a material can withstand without the occurrence of breakdown is called its **dielectric strength.** This quantity is affected significantly by temperature, trace impurities, small irregularities in the metal electrodes, and other factors that are difficult to control. For this reason we can give only approximate figures for dielectric strengths. The dielectric strength of dry air is about $3 \times 10^6\ \text{V/m}$. **Table 24.2** lists the dielectric strengths of a few common insulating materials. All of the values are substantially greater than the value for air. For example, a layer of polycarbonate 0.01 mm thick (about the smallest practical thickness) has 10 times the dielectric strength of air and can withstand a maximum voltage of about $(3 \times 10^7\ \text{V/m})(1 \times 10^{-5}\ \text{m}) = 300\ \text{V}$.

BIO APPLICATION Dielectric Cell Membrane The membrane of a living cell behaves like a dielectric between the plates of a capacitor. The membrane is made of two sheets of lipid molecules, with their water-insoluble ends in the middle and their water-soluble ends (shown in red) on the outer surfaces. Conductive fluids on either side of the membrane (water with negative ions inside the cell, water with positive ions outside) act as charged capacitor plates, and the nonconducting membrane acts as a dielectric with K of about 10. The potential difference V across the membrane is about 0.07 V and the membrane thickness d is about 7×10^{-9} m, so the electric field $E = V/d$ in the membrane is about 10^7 V/m—close to the dielectric strength of the membrane. If the membrane were made of air, V and E would be larger by a factor of $K \approx 10$ and dielectric breakdown would occur.

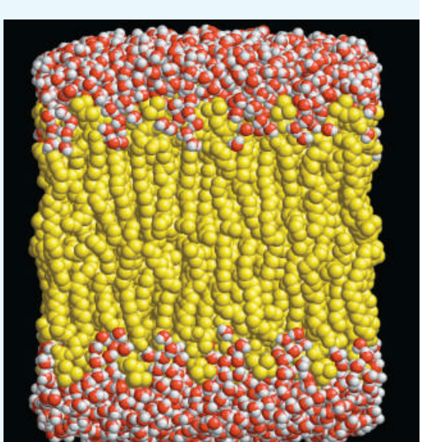


TABLE 24.2 Dielectric Constant and Dielectric Strength of Some Insulating Materials

Material	Dielectric Constant, K	Dielectric Strength, E_m (V/m)
Polycarbonate	2.8	3×10^7
Polyester	3.3	6×10^7
Polypropylene	2.2	7×10^7
Polystyrene	2.6	2×10^7
Pyrex® glass	4.7	1×10^7

TEST YOUR UNDERSTANDING OF SECTION 24.4 The space between the plates of an isolated parallel-plate capacitor is filled by a slab of dielectric with dielectric constant K. The two plates of the capacitor have charges Q and $-Q$. You pull out the dielectric slab. If the charges do not change, how does the energy in the capacitor change when you remove the slab? (i) It increases; (ii) it decreases; (iii) it remains the same.

ANSWER

(i) Here Q remains the same, so we use $U = Q^2/2C$ from Eq. (24.9) for the stored energy. Removing the dielectric lowers the capacitance by a factor of $1/K$; since U is inversely proportional to C, the stored energy *increases* by a factor of K. It takes work to pull the dielectric slab out of the capacitor because the fringing field tries to pull the slab back in (Fig. 24.16). The work that you do goes into the energy stored in the capacitor.

24.5 MOLECULAR MODEL OF INDUCED CHARGE

In Section 24.4 we discussed induced surface charges on a dielectric in an electric field. Now let's look at how these surface charges can arise. If the material were a *conductor,* the answer would be simple. Conductors contain charge that is free to move, and when an electric field is present, some of the charge redistributes itself on the surface so that there is no electric field inside the conductor. But an ideal dielectric has *no* charges that are free to move, so how can a surface charge occur?

To understand this, we have to look again at rearrangement of charge at the *molecular* level. Some molecules, such as H_2O and N_2O, have equal amounts of positive and negative charges but a lopsided distribution, with excess positive charge concentrated on one side of the molecule and negative charge on the other. As we described in Section 21.7, such an arrangement is called an *electric dipole,* and the molecule is called a *polar molecule.* When no electric field is present in a gas or liquid with polar molecules, the molecules are oriented randomly (**Fig. 24.17a**). In an electric field, however, they tend to orient themselves as in Fig. 24.17b, as a result of the electric-field torques described in Section 21.7. Because of thermal agitation, the alignment of the molecules with $\vec{E}$ is not perfect.

Even a molecule that is *not* ordinarily polar *becomes* a dipole when it is placed in an electric field because the field pushes the positive charges in the molecules in the direction of the field and pushes the negative charges in the opposite direction. This causes a redistribution of charge within the molecule (**Fig. 24.18**). Such dipoles are called *induced* dipoles.

With either polar or nonpolar molecules, the redistribution of charge caused by the field leads to the formation of a layer of charge on each surface of the dielectric material (**Fig. 24.19**). These layers are the surface charges described in Section 24.4; their surface charge density is denoted by σ_i. The charges are *not* free to move indefinitely, as they

Figure 24.17 Polar molecules (a) without and (b) with an applied electric field $\vec{E}$.

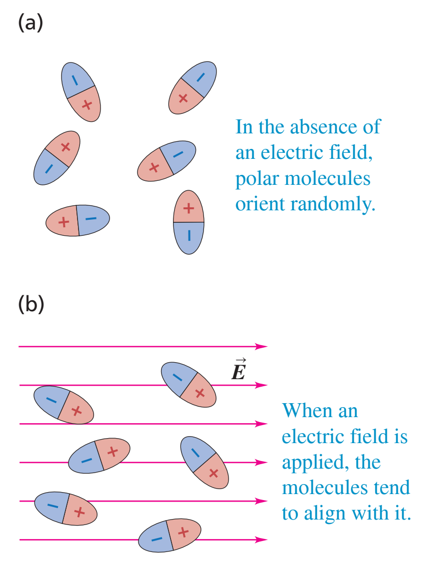


Figure 24.18 Nonpolar molecules (a) without and (b) with an applied electric field $\vec{E}$.

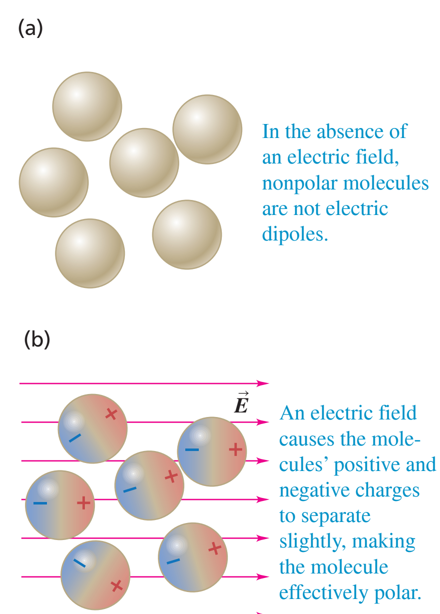
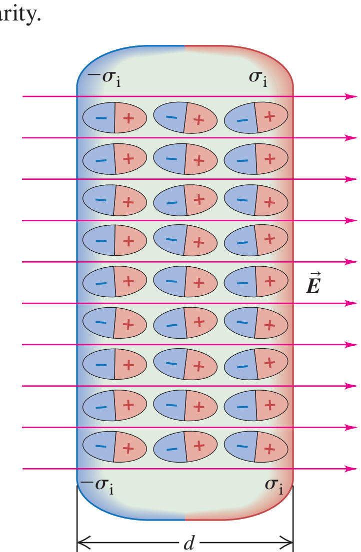
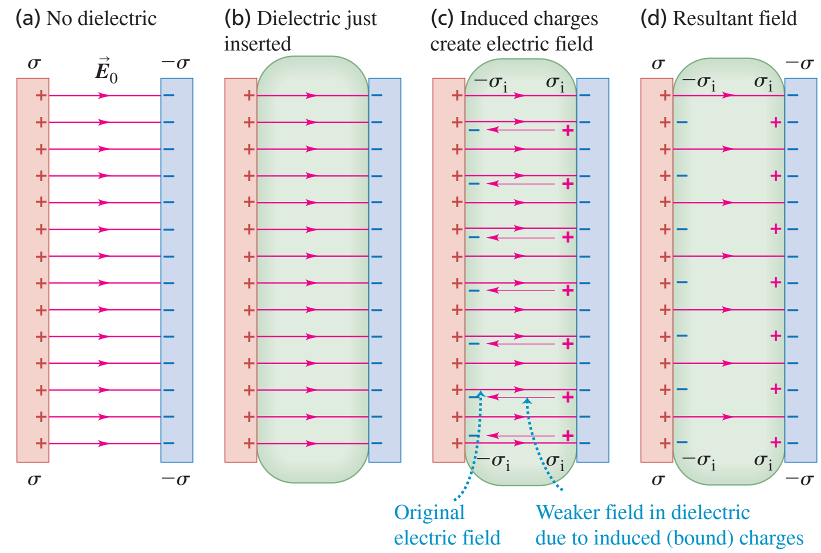


Figure 24.19 Polarization of a dielectric in an electric field $\vec{E}$ gives rise to thin layers of bound charges on the surfaces, creating surface charge densities σ_i and $-\sigma_i$. The sizes of the molecules are greatly exaggerated for clarity.



APPLICATION Smartphones, Capacitors, and Dielectrics A typical smartphone contains several hundred miniature capacitors, each of which uses a ceramic dielectric between its plates to boost its capacitance. These capacitors are used in every subsystem of the smartphone, including the main processor, the dynamic random access memory, the on–off and volume controls, the microphone, and the touch screen (see page 787).

Figure 24.20 (a) Electric field of magnitude E_0 between two charged plates. (b) Introduction of a dielectric of dielectric constant K. (c) The induced surface charges and their field. (d) Resultant field of magnitude E_0/K.

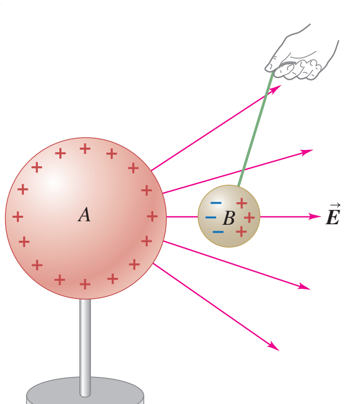


would be in a conductor, because each charge is bound to a molecule. They are in fact called **bound charges** to distinguish them from the **free charges** that are added to and removed from the conducting capacitor plates. In the interior of the material the net charge per unit volume remains zero. As we have seen, this redistribution of charge is called *polarization,* and we say that the material is *polarized*.

The four parts of **Fig. 24.20** show the behavior of a slab of dielectric when it is inserted in the field between a pair of oppositely charged capacitor plates. Figure 24.20a shows the original field. Figure 24.20b is the situation after the dielectric has been inserted but before any rearrangement of charges has occurred. Figure 24.20c shows by thinner arrows the additional field set up in the dielectric by its induced surface charges. This field is *opposite* to the original field, but it is not great enough to cancel the original field completely because the charges in the dielectric are not free to move indefinitely. The resultant field in the dielectric, shown in Fig. 24.20d, is therefore decreased in magnitude. In the field-line representation, some of the field lines leaving the positive plate go through the dielectric, while others terminate on the induced charges on the faces of the dielectric.

As we discussed in Section 21.2, polarization is also the reason a charged object, such as an electrified plastic rod, can exert a force on an *uncharged* object such as a bit of paper or a pith ball. **Figure 24.21** shows an uncharged dielectric sphere B in the radial field of a positively charged object A. The induced positive charges on B experience a force toward the right, while the force on the induced negative charges is toward the left. The negative charges are closer to A, and thus are in a stronger field, than are the positive charges. The force toward the left is stronger than that toward the right, and B is attracted toward A, even though its net charge is zero. The attraction occurs whether the sign of A's charge is positive or negative (see Fig. 21.8). Furthermore, the effect is not limited to dielectrics; an uncharged conducting object would be attracted in the same way.

Figure 24.21 A neutral sphere B in the radial electric field of a positively charged sphere A is attracted to the charge because of polarization.



TEST YOUR UNDERSTANDING OF SECTION 24.5 A parallel-plate capacitor has charges Q and $-Q$ on its two plates. A dielectric slab with $K = 3$ is then inserted into the space between the plates as shown in Fig. 24.20. Rank the following electric-field magnitudes in order from largest to smallest. (i) The field before the slab is inserted; (ii) the resultant field after the slab is inserted; (iii) the field due to the bound charges.

ANSWER

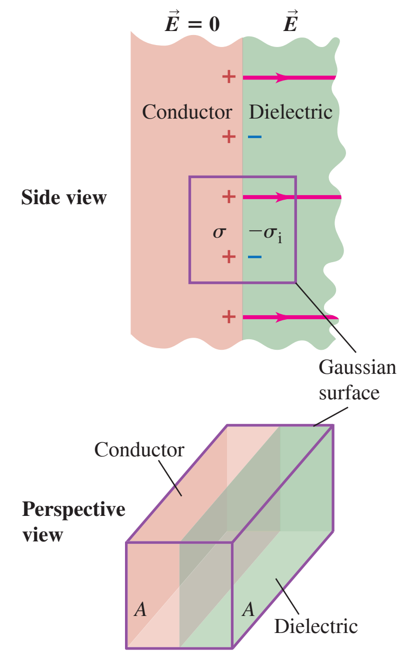
(i), (iii), (ii) Equation (24.14) says that if E_0 is the initial electric-field magnitude (before the dielectric slab is inserted), then the resultant field magnitude after the slab is inserted is $E_0/K = E_0/3$. The magnitude of the resultant field equals the difference between the initial field magnitude and the magnitude E_i of the field due to the bound charges (see Fig. 24.20). Hence $E_0 - E_i = E_0/3$ and $E_i = 2E_0/3$.

24.6 GAUSS'S LAW IN DIELECTRICS

We can extend the analysis of Section 24.4 to reformulate Gauss's law in a form that is particularly useful for dielectrics. **Figure 24.22** is a close-up view of the left capacitor plate and left surface of the dielectric in Fig. 24.15b. Let's apply Gauss's law to the rectangular box shown in cross section by the purple line; the surface area of the left and right sides is A. The left side is embedded in the conductor that forms the left capacitor plate, and so the electric field everywhere on that surface is zero. The right side is embedded in the dielectric, where the electric field has magnitude E, and $E_\perp = 0$ everywhere on the other four sides. The total charge enclosed, including both the charge on the capacitor plate and the induced charge on the dielectric surface, is $Q_{\text{encl}} = (\sigma - \sigma_{\text{i}})A$, so Gauss's law gives

$$EA = \frac{(\sigma - \sigma_{\text{i}})A}{\epsilon_0} \qquad (24.21)$$

Figure **24.22** Gauss's law with a dielectric. This figure shows a close-up of the left-hand capacitor plate in Fig. 24.15b. The Gaussian surface is a rectangular box that lies half in the conductor and half in the dielectric.



This equation is not very illuminating as it stands because it relates two unknown quantities: E inside the dielectric and the induced surface charge density σ_{i}. But now we can use Eq. (24.16), developed for this same situation, to simplify this equation by eliminating σ_{i}. Equation (24.16) is

$$\sigma_{\text{i}} = \sigma\left(1 - \frac{1}{K}\right)$$

or

$$\sigma - \sigma_{\text{i}} = \frac{\sigma}{K}$$

Combining this with Eq. (24.21), we get

$$EA = \frac{\sigma A}{K\epsilon_0}$$

or

$$KEA = \frac{\sigma A}{\epsilon_0} \qquad (24.22)$$

Equation (24.22) says that the flux of $K\vec{E}$, not $\vec{E}$, through the Gaussian surface in Fig. 24.22 is equal to the enclosed *free* charge σA divided by ϵ_0. It turns out that for *any* Gaussian surface, whenever the induced charge is proportional to the electric field in the material, we can rewrite Gauss's law as

Gauss's law in a dielectric:

Dielectric constant

$$\oint K\vec{E} \cdot d\vec{A} = \frac{Q_{\text{encl-free}}}{\epsilon_0} \qquad (24.23)$$

Total free charge enclosed by surface

Surface integral of $K\vec{E}$ over a closed surface

Electric constant

where $Q_{\text{encl-free}}$ is the total *free* charge (not bound charge) enclosed by the Gaussian surface. The significance of these results is that the right sides contain only the *free* charge on the conductor, not the bound (induced) charge. In fact, although we have not proved it, Eq. (24.23) remains valid even when different parts of the Gaussian surface are embedded in dielectrics having different values of K, provided that the value of K in each dielectric is independent of the electric field (usually the case for electric fields that are not too strong) and that we use the appropriate value of K for each point on the Gaussian surface.

EXAMPLE 24.12 A spherical capacitor with dielectric

Use Gauss's law to find the capacitance of the spherical capacitor of Example 24.3 (Section 24.1) if the volume between the shells is filled with an insulating oil with dielectric constant K.

IDENTIFY and SET UP The spherical symmetry of the problem is not changed by the presence of the dielectric, so as in Example 24.3, we use a concentric spherical Gaussian surface of radius r between the shells. Since a dielectric is present, we use Gauss's law in the form of Eq. (24.23).

EXECUTE From Eq. (24.23),

$$\oint K\vec{E} \cdot d\vec{A} = \oint KE\,dA = KE \oint dA = (KE)(4\pi r^2) = \frac{Q}{\epsilon_0}$$

$$E = \frac{Q}{4\pi K\epsilon_0 r^2} = \frac{Q}{4\pi\epsilon r^2}$$

where $\epsilon = K\epsilon_0$. Compared to the case in which there is vacuum between the shells, the electric field is reduced by a factor of $1/K$. The potential difference V_{ab} between the shells is reduced by the same factor, and so the capacitance $C = Q/V_{ab}$ is *increased* by a factor of K, just as for a parallel-plate capacitor when a dielectric is inserted. Using the result of Example 24.3, we find that the capacitance with the dielectric is

$$C = \frac{4\pi K\epsilon_0 r_a r_b}{r_b - r_a} = \frac{4\pi\epsilon r_a r_b}{r_b - r_a}$$

EVALUATE If the dielectric fills the volume between the two conductors, the capacitance is just K times the value with no dielectric. The result is more complicated if the dielectric only partially fills this volume.

KEYCONCEPT Gauss's law can be written in a modified way if a dielectric is present. If the volume of space being studied is filled with a dielectric that has dielectric constant K, then the flux of the *product* of K and the electric field through a closed surface is proportional to the amount of *free* charge enclosed by the surface.

TEST YOUR UNDERSTANDING OF SECTION 24.6 A single point charge q is embedded in a very large block of dielectric of dielectric constant K. At a point inside the dielectric a distance r from the point charge, what is the magnitude of the electric field? (i) $q/4\pi\epsilon_0 r^2$; (ii) $Kq/4\pi\epsilon_0 r^2$; (iii) $q/4\pi K\epsilon_0 r^2$; (iv) none of these.

ANSWER

(iii) Equation (24.23) shows that this situation is the same as an isolated point charge in vacuum but with $\vec{E}$ replaced by $K\vec{E}$. Hence KE at the point of interest is equal to $q/4\pi\epsilon_0 r^2$, and so $E = q/4\pi K\epsilon_0 r^2$. As in Example 24.12, filling the space with a dielectric reduces the electric field by a factor of $1/K$.

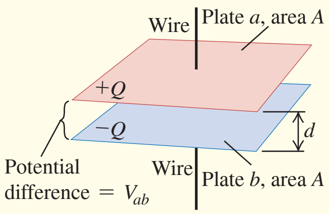
CHAPTER 24 SUMMARY

Capacitors and capacitance: A capacitor is any pair of conductors separated by an insulating material. When the capacitor is charged, there are charges of equal magnitude Q and opposite sign on the two conductors, and the potential V_{ab} of the positively charged conductor with respect to the negatively charged conductor is proportional to Q. The capacitance C is defined as the ratio of Q to V_{ab}. The SI unit of capacitance is the farad (F): $1\text{ F} = 1\text{ C/V}$.

A parallel-plate capacitor consists of two parallel conducting plates, each with area A, separated by a distance d. If they are separated by vacuum, the capacitance depends on only A and d. For other geometries, the capacitance can be found by using the definition $C = Q/V_{ab}$. (See Examples 24.1–24.4.)

$$C = \frac{Q}{V_{ab}} \qquad (24.1)$$

$$C = \frac{Q}{V_{ab}} = \epsilon_0 \frac{A}{d} \qquad (24.2)$$



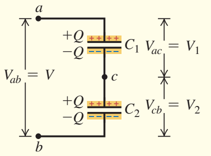
Capacitors in series and parallel: When capacitors with capacitances $C_1, C_2, C_3, \ldots$ are connected in series, the reciprocal of the equivalent capacitance C_{eq} equals the sum of the reciprocals of the individual capacitances. When capacitors are connected in parallel, the equivalent capacitance C_{eq} equals the sum of the individual capacitances. (See Examples 24.5 and 24.6.)

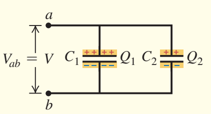
$$\frac{1}{C_{eq}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \cdots \qquad (24.5)$$

(capacitors in series)

$$C_{eq} = C_1 + C_2 + C_3 + \cdots \qquad (24.7)$$

(capacitors in parallel)

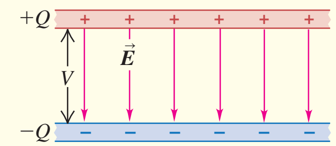




Energy in a capacitor: The energy U required to charge a capacitor C to a potential difference V and a charge Q is equal to the energy stored in the capacitor. This energy can be thought of as residing in the electric field between the conductors; the energy density u (energy per unit volume) is proportional to the square of the electric-field magnitude. (See Examples 24.7–24.9.)

$$U = \frac{Q^2}{2C} = \tfrac{1}{2}CV^2 = \tfrac{1}{2}QV \qquad (24.9)$$

$$u = \tfrac{1}{2}\epsilon_0 E^2 \qquad (24.11)$$



Dielectrics: When the space between the conductors is filled with a dielectric material, the capacitance increases by a factor K, the dielectric constant of the material. The quantity $\epsilon = K\epsilon_0$ is the permittivity of the dielectric. For a fixed amount of charge on the capacitor plates, induced charges on the surface of the dielectric decrease the electric field and potential difference between the plates by the same factor K. The surface charge results from polarization, a microscopic rearrangement of charge in the dielectric. (See Example 24.10.)

Under sufficiently strong electric fields, dielectrics become conductors, a situation called dielectric breakdown. The maximum field that a material can withstand without breakdown is called its dielectric strength.

In a dielectric, the expression for the energy density is the same as in vacuum but with ϵ_0 replaced by $\epsilon = K\epsilon_0$. (See Example 24.11.)

Gauss's law in a dielectric has almost the same form as in vacuum, with two key differences: $\vec{E}$ is replaced by $K\vec{E}$ and Q_{encl} is replaced by $Q_{\text{encl-free}}$, which includes only the free charge (not bound charge) enclosed by the Gaussian surface. (See Example 24.12.)

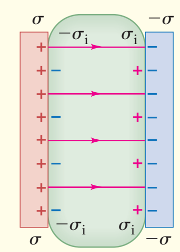
$$C = KC_0 = K\epsilon_0\frac{A}{d} = \epsilon\frac{A}{d} \qquad (24.19)$$

(parallel-plate capacitor filled with dielectric)

$$u = \tfrac{1}{2}K\epsilon_0 E^2 = \tfrac{1}{2}\epsilon E^2 \qquad (24.20)$$

$$\oint K\vec{E} \cdot d\vec{A} = \frac{Q_{\text{encl-free}}}{\epsilon_0} \qquad (24.23)$$

Dielectric between plates



Chapter 24 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 24.2, 24.3, and 24.4 (Section 24.1) before attempting these problems.

VP24.4.1 The plates of a parallel-plate capacitor in vacuum are 3.70 mm apart and 2.75 m^2 in area. When you apply a certain potential difference across the capacitor, the surface charge density on the positive plate is 1.40×10^{-5} C/m^2. Calculate (a) the capacitance of the capacitor and (b) the potential difference.

VP24.4.2 One model of a parallel-plate capacitor in vacuum is advertised as having a capacitance of 8.30 nF. When you disassemble one such capacitor, you find that its plates are 4.50 m^2 in area. (a) What is the spacing between this capacitor's plates? (b) When the potential difference between the plates is 3.00 kV, what are the charge on the capacitor and the magnitude of the electric field between the plates?

VP24.4.3 A spherical capacitor has vacuum between its conducting shells and a capacitance of 125 pF. The outer shell has inner radius 9.00 cm. (a) What is the outer radius of the inner shell? If the potential of the inner shell is 355 V higher than the potential of the outer shell, what are (b) the capacitor charge and (c) the surface charge densities on the inner and outer shells?

VP24.4.4 A cylindrical capacitor has a capacitance per length of 69.0 pF/m. (a) What is the ratio of the inner radius of the outer cylinder to the outer radius of the inner cylinder? (b) If the outer conductor carries charge –8.62 nC per meter of length, what is the potential difference between the two conductors? Which conductor is at higher potential?

Be sure to review EXAMPLES 24.5 and 24.6 (Section 24.2) and EXAMPLES 24.7, 24.8, and 24.9 (Section 24.3) before attempting these problems.

VP24.9.1 You have three capacitors that you wish to arrange in a circuit: $C_1 = 1.00\ \mu$F, $C_2 = 2.50\ \mu$F, and $C_3 = 5.00\ \mu$F. Find their equivalent capacitance if (a) they are connected in series; (b) they are connected in parallel; (c) C_1 and C_2 are connected in parallel, and that combination is connected to C_3 in series; (d) C_1 and C_2 are connected in series, and that combination is connected to C_3 in parallel.

VP24.9.2 You have two capacitors that you wish to use in an energy-storage device: $C_1 = 1.50\ \mu$F and $C_2 = 6.00\ \mu$F. (a) How much energy is stored in each capacitor if each has charge 4.50×10^{-4} C? Which has greater stored energy? (b) How much energy is stored in each capacitor if each has potential difference 125 V? Which has greater stored energy?

VP24.9.3 You connect two capacitors, $C_1 = 1.50\ \mu$F and $C_2 = 4.00\ \mu$F. (a) How much energy is stored in the combination if the potential difference across the combination is 145 V and the capacitors are connected in series? In parallel? In which case is the stored energy greater? (b) How much energy is stored in the combination if the charge on the combination is 4.00×10^{-4} C and the capacitors are connected in series? In parallel? In which case is the stored energy greater?

VP24.9.4 The plates of a parallel-plate capacitor in vacuum are 1.40 mm apart and 2.45 m^2 in area. The charge on the capacitor is 1.25×10^{-5} C. Calculate (a) the surface charge density on the positive plate, (b) the electric field magnitude between the plates, (c) the electric energy density between the plates, and (d) the total energy stored in the field between the plates.

Be sure to review EXAMPLES 24.10 and 24.11 (Section 24.4) before attempting these problems.

VP24.11.1 The two plates of a parallel-plate capacitor each have area $0.460\ \text{m}^2$, are 3.00 mm apart, and initially have vacuum between them. A power supply is attached to the capacitor, charging it to 4.00 kV, and is then disconnected. A dielectric sheet is then inserted that fills the space between the plates. The potential difference between the plates decreases to 2.50 kV, and the charge on each plate remains constant. Find (a) the original capacitance, (b) the charge on the capacitor plates, (c) the dielectric constant of the dielectric, (d) the new capacitance after you insert the dielectric, and (e) the magnitude of the induced charge on the faces of the dielectric sheet.

VP24.11.2 The two plates of a parallel-plate capacitor each have area $1.20\ \text{m}^2$, are 4.50 mm apart, and initially have vacuum between them. You connect a power supply to the capacitor, charging it to 3.50 kV. Keeping the power supply connected, you insert a sheet of dielectric with dielectric constant 2.50 that fills the space between the plates. Find (a) the original capacitance, (b) the original charge on the capacitor plates, (c) the new capacitance once you insert the dielectric, (d) the new charge on the capacitor plates, and (e) the magnitude of the induced charge on the faces of the dielectric sheet.

VP24.11.3 With vacuum between its plates, a parallel-plate capacitor has capacitance $4.50\ \mu\text{F}$. You attach a power supply to the capacitor, charging it to 2.80 kV, and then disconnect it. You then insert a dielectric sheet that fills the space between the plates. The potential difference between the plates decreases to 1.20 kV, and the charge on each plate remains constant. Find the energy stored in the capacitor (a) before you insert the sheet and (b) after you insert the sheet.

VP24.11.4 With vacuum between its plates, a parallel-plate capacitor has capacitance $5.60\ \mu\text{F}$. You attach a power supply to the capacitor, charging it to 1.50 kV. Keeping the power supply connected, you insert a dielectric sheet with dielectric constant 2.85 that fills the space between the plates. Find the energy stored in the capacitor (a) before the sheet is inserted and (b) after the sheet is inserted.

BRIDGING PROBLEM Electric-Field Energy and Capacitance of a Conducting Sphere

A solid conducting sphere of radius R carries a charge Q. Calculate the electric-field energy density at a point a distance r from the center of the sphere for (a) $r < R$ and (b) $r > R$. (c) Calculate the total electric-field energy associated with the charged sphere. (d) How much work is required to assemble the charge Q on the sphere? (e) Use the result of part (c) to find the capacitance of the sphere. (You can think of the second conductor as a hollow conducting shell of infinite radius.)

SOLUTION GUIDE

IDENTIFY and SET UP

1. You know the electric field for this situation at all values of r from Example 22.5 (Section 22.4). You'll use this to find the electric-field energy density u and *total* electric-field energy U. You can then find the capacitance from the relationship $U = Q^2/2C$.
2. To find U, consider a spherical shell of radius r and thickness dr that has volume $dV = 4\pi r^2 dr$. (It will help to make a drawing of such a shell concentric with the conducting sphere.) The energy stored in this volume is $u\,dV$, and the total energy is the integral of $u\,dV$ from $r = 0$ to $r \rightarrow \infty$. Set up this integral.

EXECUTE

3. Find u for $r < R$ and for $r > R$. (*Hint:* What is the field inside a solid conductor?)
4. Substitute your results from step 3 into the expression from step 2. Then calculate the integral to find the total electric-field energy U.
5. Use your understanding of the energy stored in a charge distribution to find the work required to assemble the charge Q.
6. Find the capacitance of the sphere.

EVALUATE

7. Where is the electric-field energy density greatest? Where is it least?
8. How would the results be affected if the solid sphere were replaced by a hollow conducting sphere of the same radius R?
9. You can find the potential difference between the sphere and infinity from $C = Q/V$. Does this agree with the result of Example 23.8 (Section 23.3)?

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q24.1 Equation (24.2) shows that the capacitance of a parallel-plate capacitor becomes larger as the plate separation d decreases. However, there is a practical limit to how small d can be made, which places limits on how large C can be. Explain what sets the limit on d. (*Hint:* What happens to the magnitude of the electric field as $d \rightarrow 0$?)

Q24.2 Suppose several different parallel-plate capacitors are charged up by a constant-voltage source. Thinking of the actual movement and position of the charges on an atomic level, why does it make sense that the capacitances are proportional to the surface areas of the plates? Why does it make sense that the capacitances are *inversely* proportional to the distance between the plates?

Q24.3 Suppose the two plates of a capacitor have different areas. When the capacitor is charged by connecting it to a battery, do the charges on the two plates have equal magnitude, or may they be different? Explain your reasoning.

Q24.4 To store the maximum amount of energy in a parallel-plate capacitor with a given battery (voltage source), would it be better to have the plates far apart or close together?

Q24.5 In the parallel-plate capacitor of Fig. 24.2, suppose the plates are pulled apart so that the separation d is much larger than the size of the plates. (a) Is it still accurate to say that the electric field between the plates is uniform? Why or why not? (b) In the situation shown in Fig. 24.2, the potential difference between the plates is $V_{ab} = Qd/\epsilon_0 A$. If the plates are pulled apart as described above, is V_{ab} more or less than this formula would indicate? Explain your reasoning. (c) With the plates pulled apart as described above, is the capacitance more than, less than, or the same as that given by Eq. (24.2)? Explain your reasoning.

Q24.6 A parallel-plate capacitor is charged by being connected to a battery and is kept connected to the battery. The separation between the plates is then doubled. How does the electric field change? The charge on the plates? The total energy? Explain.

Q24.7 A parallel-plate capacitor is charged by being connected to a battery and is then disconnected from the battery. The separation between the plates is then doubled. How does the electric field change? The potential difference? The total energy? Explain.

Q24.8 Two parallel-plate capacitors, identical except that one has twice the plate separation of the other, are charged by the same voltage source. Which capacitor has a stronger electric field between the plates? Which capacitor has a greater charge? Which has greater energy density? Explain your reasoning.

Q24.9 The charged plates of a capacitor attract each other, so to pull the plates farther apart requires work by some external force. What becomes of the energy added by this work? Explain.

Q24.10 You have two capacitors and want to connect them across a voltage source (battery) to store the maximum amount of energy. Should they be connected in series or in parallel?

Q24.11 As shown in Table 24.1, water has a very large dielectric constant $K = 80.4$. Why do you think water is not commonly used as a dielectric in capacitors?

Q24.12 Is dielectric strength the same thing as dielectric constant? Explain any differences between the two quantities. Is there a simple relationship between dielectric strength and dielectric constant (see Table 24.2)?

Q24.13 A capacitor made of aluminum foil strips separated by Mylar film was subjected to excessive voltage, and the resulting dielectric breakdown melted holes in the Mylar. After this, the capacitance was found to be about the same as before, but the breakdown voltage was much less. Why?

Q24.14 Suppose you bring a slab of dielectric close to the gap between the plates of a charged capacitor, preparing to slide it between the plates. What force will you feel? What does this force tell you about the energy stored between the plates once the dielectric is in place, compared to before the dielectric is in place?

Q24.15 The freshness of fish can be measured by placing a fish between the plates of a capacitor and measuring the capacitance. How does this work? (*Hint:* As time passes, the fish dries out. See Table 24.1.)

Q24.16 *Electrolytic* capacitors use as their dielectric an extremely thin layer of nonconducting oxide between a metal plate and a conducting solution. Discuss the advantage of such a capacitor over one constructed using a solid dielectric between the metal plates.

Q24.17 In terms of the dielectric constant K, what happens to the electric flux through the Gaussian surface shown in Fig. 24.22 when the dielectric is inserted into the previously empty space between the plates? Explain.

Q24.18 A parallel-plate capacitor is connected to a power supply that maintains a fixed potential difference between the plates. (a) If a sheet of dielectric is then slid between the plates, what happens to (i) the electric field between the plates, (ii) the magnitude of charge on each plate, and (iii) the energy stored in the capacitor? (b) Now suppose that before the dielectric is inserted, the charged capacitor is disconnected from the power supply. In this case, what happens to (i) the electric field between the plates, (ii) the magnitude of charge on each plate, and (iii) the energy stored in the capacitor? Explain any differences between the two situations.

Q24.19 Liquid dielectrics that have polar molecules (such as water) always have dielectric constants that decrease with increasing temperature. Why?

Q24.20 A conductor is an extreme case of a dielectric, since if an electric field is applied to a conductor, charges are free to move within the conductor to set up "induced charges." What is the dielectric constant of a perfect conductor? Is it $K = 0$, $K \rightarrow \infty$, or something in between? Explain your reasoning.

Q24.21 The two plates of a capacitor are given charges $\pm Q$. The capacitor is then disconnected from the charging device so that the charges on the plates can't change, and the capacitor is immersed in a tank of oil. Does the electric field between the plates increase, decrease, or stay the same? Explain your reasoning. How can this field be measured?

EXERCISES

Section 24.1 Capacitors and Capacitance

24.1 • The plates of a parallel-plate capacitor are 2.50 mm apart, and each carries a charge of magnitude 80.0 nC. The plates are in vacuum. The electric field between the plates has a magnitude of 4.00×10^6 V/m. What is (a) the potential difference between the plates; (b) the area of each plate; (c) the capacitance?

24.2 • The plates of a parallel-plate capacitor are 3.28 mm apart, and each has an area of 9.82 cm^2. Each plate carries a charge of magnitude 4.35×10^{-8} C. The plates are in vacuum. What is (a) the capacitance; (b) the potential difference between the plates; (c) the magnitude of the electric field between the plates?

24.3 • A parallel-plate air capacitor of capacitance 245 pF has a charge of magnitude 0.148 μC on each plate. The plates are 0.328 mm apart. (a) What is the potential difference between the plates? (b) What is the area of each plate? (c) What is the electric-field magnitude between the plates? (d) What is the surface charge density on each plate?

24.4 • A 5.00 μF parallel-plate capacitor is connected to a 12.0 V battery. After the capacitor is fully charged, the battery is disconnected without loss of any of the charge on the plates. (a) A voltmeter is connected across the two plates without discharging them. What does it read? (b) What would the voltmeter read if (i) the plate separation were doubled; (ii) the radius of each plate were doubled but their separation was unchanged?

24.5 • A 10.0 μF parallel-plate capacitor with circular plates is connected to a 12.0 V battery. (a) What is the charge on each plate? (b) How much charge would be on the plates if their separation were doubled while the capacitor remained connected to the battery? (c) How much charge would be on the plates if the capacitor were connected to the 12.0 V battery after the radius of each plate was doubled without changing their separation?

24.6 • A 5.00 pF, parallel-plate, air-filled capacitor with circular plates is to be used in a circuit in which it will be subjected to potentials of up to 1.00×10^2 V. The electric field between the plates is to be no greater than 1.00×10^4 N/C. As a budding electrical engineer for Live-Wire Electronics, your tasks are to (a) design the capacitor by finding what its physical dimensions and separation must be; (b) find the maximum charge these plates can hold.

24.7 • A parallel-plate air capacitor is to store charge of magnitude 240.0 pC on each plate when the potential difference between the plates is 42.0 V. (a) If the area of each plate is 6.80 cm^2, what is the separation between the plates? (b) If the separation between the two plates is double the value calculated in part (a), what potential difference is required for the capacitor to store charge of magnitude 240.0 pC on each plate?

24.8 • A cylindrical capacitor consists of a solid inner conducting core with radius 0.250 cm, surrounded by an outer hollow conducting tube. The two conductors are separated by air, and the length of the cylinder is 12.0 cm. The capacitance is 36.7 pF. (a) Calculate the inner radius of the hollow tube. (b) When the capacitor is charged to 125 V, what is the charge per unit length λ on the capacitor?

24.9 • A capacitor is made from two hollow, coaxial, iron cylinders, one inside the other. The inner cylinder is negatively charged and the outer is positively charged; the magnitude of the charge on each is 10.0 pC. The inner cylinder has radius 0.50 mm, the outer one has radius 5.00 mm, and the length of each cylinder is 18.0 cm. (a) What is the capacitance? (b) What applied potential difference is necessary to produce these charges on the cylinders?

24.10 •• A cylindrical capacitor has an inner conductor of radius 2.2 mm and an outer conductor of radius 3.5 mm. The two conductors are separated by vacuum, and the entire capacitor is 2.8 m long. (a) What is the capacitance per unit length? (b) The potential of the inner conductor is 350 mV higher than that of the outer conductor. Find the charge (magnitude and sign) on both conductors.

24.11 •• A spherical capacitor contains a charge of 3.30 nC when connected to a potential difference of 220 V. If its plates are separated by vacuum and the inner radius of the outer shell is 4.00 cm, calculate: (a) the capacitance; (b) the radius of the inner sphere; (c) the electric field just outside the surface of the inner sphere.

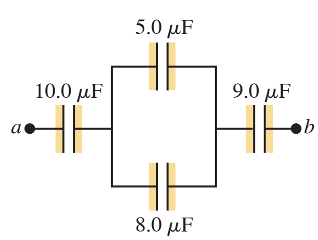
24.12 • A spherical capacitor is formed from two concentric, spherical, conducting shells separated by vacuum. The inner sphere has radius 15.0 cm and the capacitance is 116 pF. (a) What is the radius of the outer sphere? (b) If the potential difference between the two spheres is 220 V, what is the magnitude of charge on each sphere?

Section 24.2 Capacitors in Series and Parallel

24.13 •• You measure the capacitance C_1 of a capacitor by doing the following: First connect capacitors C_1 and C_2 in series to a power supply that provides a voltage V that can be varied. The capacitance of C_2 is known to be 3.00 μF. Then vary the applied voltage V, and for each value of V measure the voltage V_2 across C_2. After plotting your data as V_2 versus V, you find that the data fall close to a straight line that has slope 0.650. What is the capacitance C_1?

24.14 • **Figure E24.14** shows a system of four capacitors, where the potential difference across ab is 50.0 V. (a) Find the equivalent capacitance of this system between a and b. (b) How much charge is stored by this combination of capacitors? (c) How much charge is stored in each of the 10.0 μF and the 9.0 μF capacitors?

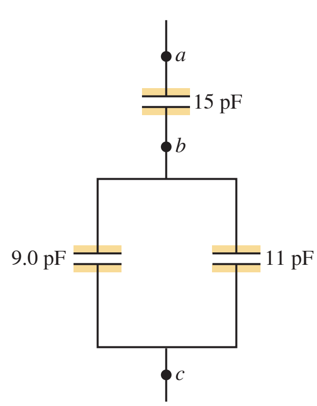
Figure **E24.14**



24.15 • **BIO** **Electric Eels.** Electric eels and electric fish generate large potential differences that are used to stun enemies and prey. These potentials are produced by cells that each can generate 0.10 V. We can plausibly model such cells as charged capacitors. (a) How should these cells be connected (in series or in parallel) to produce a total potential of more than 0.10 V? (b) Using the connection in part (a), how many cells must be connected together to produce the 500 V surge of the electric eel?

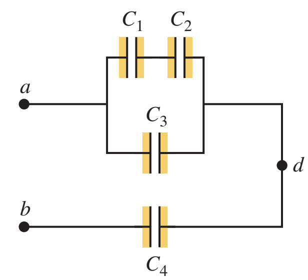
24.16 • For the system of capacitors shown in **Fig. E24.16**, find the equivalent capacitance (a) between b and c, and (b) between a and c.

Figure **E24.16**



24.17 • In **Fig. E24.17**, each capacitor has $C = 4.00\ \mu$F and $V_{ab} = +28.0$ V. Calculate (a) the charge on each capacitor; (b) the potential difference across each capacitor; (c) the potential difference between points a and d.

Figure **E24.17**

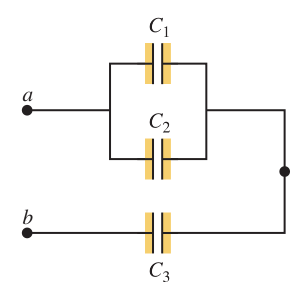


24.18 • In Fig. 24.8a, let $C_1 = 3.00\ \mu$F, $C_2 = 5.00\ \mu$F, and $V_{ab} = +64.0$ V. Calculate (a) the charge on each capacitor and (b) the potential difference across each capacitor.

24.19 • In Fig. 24.9a, let $C_1 = 3.00\ \mu$F, $C_2 = 5.00\ \mu$F, and $V_{ab} = +52.0$ V. Calculate (a) the charge on each capacitor and (b) the potential difference across each capacitor.

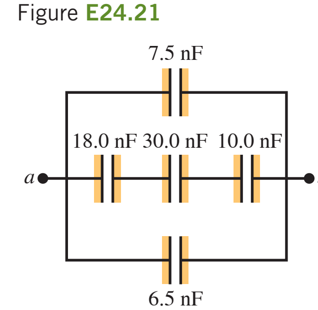
24.20 • In **Fig. E24.20**, $C_1 = 6.00\ \mu$F, $C_2 = 3.00\ \mu$F, and $C_3 = 5.00\ \mu$F. The capacitor network is connected to an applied potential V_{ab}. After the charges on the capacitors have reached their final values, the charge on C_2 is 30.0 μC. (a) What are the charges on capacitors C_1 and C_3? (b) What is the applied voltage V_{ab}?

Figure **E24.20**



24.21 •• For the system of capacitors shown in **Fig. E24.21**, a potential difference of 25 V is maintained across ab. (a) What is the equivalent capacitance of this system between a and b? (b) How much charge is stored by this system? (c) How much charge does the 6.5 nF capacitor store? (d) What is the potential difference across the 7.5 nF capacitor?

Figure **E24.21**



24.22 •• A capacitor is formed from two concentric spherical conducting shells separated by vacuum. The inner sphere has radius 12.5 cm, and the outer sphere has radius 14.8 cm. A potential difference of 120 V is applied to the capacitor. (a) What is the energy density at $r = 12.6$ cm, just outside the inner sphere? (b) What is the energy density at $r = 14.7$ cm, just inside the outer sphere? (c) For a parallel-plate capacitor the energy density is uniform in the region between the plates, except near the edges of the plates. Is this also true for a spherical capacitor?

Section 24.3 Energy Storage in Capacitors and Electric-Field Energy

24.23 • A 5.80 μF, parallel-plate, air capacitor has a plate separation of 5.00 mm and is charged to a potential difference of 400 V. Calculate the energy density in the region between the plates, in units of J/m^3.

24.24 • A parallel-plate air capacitor has a capacitance of 920 pF. The charge on each plate is 3.90 μC. (a) What is the potential difference between the plates? (b) If the charge is kept constant, what will be the potential difference if the plate separation is doubled? (c) How much work is required to double the separation?

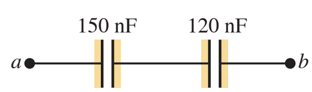
24.25 • An air capacitor is made from two flat parallel plates 1.50 mm apart. The magnitude of charge on each plate is 0.0180 μC when the potential difference is 200 V. (a) What is the capacitance? (b) What is the area of each plate? (c) What maximum voltage can be applied without dielectric breakdown? (Dielectric breakdown for air occurs at an electric-field strength of 3.0×10^6 V/m.) (d) When the charge is 0.0180 μC, what total energy is stored?

24.26 •• A parallel-plate vacuum capacitor has 8.38 J of energy stored in it. The separation between the plates is 2.30 mm. If the separation is decreased to 1.15 mm, what is the energy stored (a) if the capacitor is disconnected from the potential source so the charge on the plates remains constant, and (b) if the capacitor remains connected to the potential source so the potential difference between the plates remains constant?

24.27 • You have two identical capacitors and an external potential source. (a) Compare the total energy stored in the capacitors when they are connected to the applied potential in series and in parallel. (b) Compare the maximum amount of charge stored in each case. (c) Energy storage in a capacitor can be limited by the maximum electric field between the plates. What is the ratio of the electric field for the series and parallel combinations?

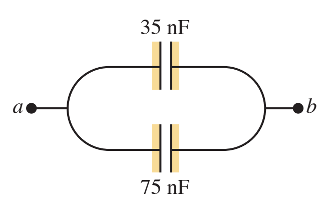
24.28 • For the capacitor network shown in **Fig. E24.28**, the potential difference across ab is 48 V. Find (a) the total charge stored in this network; (b) the charge on each capacitor; (c) the total energy stored in the network; (d) the energy stored in each capacitor; (e) the potential differences across each capacitor.

Figure **E24.28**



24.29 • For the capacitor network shown in **Fig. E24.29**, the potential difference across ab is 220 V. Find (a) the total charge stored in this network; (b) the charge on each capacitor; (c) the total energy stored in the network; (d) the energy stored in each capacitor; (e) the potential difference across each capacitor.

Figure **E24.29**



Section 24.4 Dielectrics

24.30 • Two air-filled parallel-plate capacitors with capacitances C_1 and C_2 are connected in parallel to a battery that has a voltage of 36.0 V; $C_1 = 4.00\ \mu$F and $C_2 = 6.00\ \mu$F. (a) What is the total positive charge stored in the two capacitors? (b) While the capacitors remain connected to the battery, a dielectric with dielectric constant 5.00 is inserted between the plates of capacitor C_1, completely filling the space between them. Then what is the total positive charge stored on the two capacitors? Does the insertion of the dielectric cause the total charge stored to increase or decrease?

24.31 • A 12.5 μF capacitor is connected to a power supply that keeps a constant potential difference of 24.0 V across the plates. A piece of material having a dielectric constant of 3.75 is placed between the plates, completely filling the space between them. (a) How much energy is stored in the capacitor before and after the dielectric is inserted? (b) By how much did the energy change during the insertion? Did it increase or decrease?

24.32 • A parallel-plate capacitor has capacitance $C_0 = 8.00$ pF when there is air between the plates. The separation between the plates is 1.50 mm. (a) What is the maximum magnitude of charge Q that can be placed on each plate if the electric field in the region between the plates is not to exceed 3.00×10^4 V/m? (b) A dielectric with $K = 2.70$ is inserted between the plates of the capacitor, completely filling the volume between the plates. Now what is the maximum magnitude of charge on each plate if the electric field between the plates is not to exceed 3.00×10^4 V/m?

24.33 • Two parallel plates have equal and opposite charges. When the space between the plates is evacuated, the electric field is $E = 3.20 \times 10^5$ V/m. When the space is filled with dielectric, the electric field is $E = 2.50 \times 10^5$ V/m. (a) What is the charge density on each surface of the dielectric? (b) What is the dielectric constant?

24.34 •• Two identical air-filled parallel-plate capacitors C_1 and C_2, each with capacitance C, are connected in series to a battery that has voltage V. While the two capacitors remain connected to the battery, a dielectric with dielectric constant $K > 1$ is inserted between the plates of one of the capacitors, completely filling the space between them. Let U_0 be the total energy stored in the two capacitors without the dielectric and U be the total energy stored after the dielectric is inserted. In terms of K, what is the ratio U/U_0? Does the total stored energy increase, decrease, or stay the same after the dielectric is inserted?

24.35 •• Two identical air-filled parallel-plate capacitors C_1 and C_2 are connected in series to a battery that has voltage V. The charge on each capacitor is Q_0. While the two capacitors remain connected to the battery, a dielectric with dielectric constant $K > 1$ is inserted between the plates of capacitor C_1, completely filling the space between them. In terms of K and Q_0, what is the charge on capacitor C_1 after the dielectric is inserted? Does the charge on C_1 increase, decrease, or stay the same?

24.36 •• Two air-filled parallel-plate capacitors with capacitances C_1 and C_2 are connected in series to a battery that has voltage V; $C_1 = 3.00\ \mu$F and $C_2 = 6.00\ \mu$F. The electric field between the plates of capacitor C_2 is E_{02}. While the two capacitors remain connected to the battery, a dielectric with dielectric constant $K = 4$ is inserted between the plates of capacitor C_1, completely filling the space between them. After the dielectric is inserted in C_1, the electric field between the plates of capacitor C_2 is E_2. (a) What is the ratio E_2/E_{02}? When the dielectric is inserted into C_1, does the electric field in C_2 increase, decrease, or remain the same? (b) Repeat the calculation in part (a) for the two capacitors connected to the battery in parallel.

24.37 • The dielectric to be used in a parallel-plate capacitor has a dielectric constant of 3.60 and a dielectric strength of 1.60×10^7 V/m. The capacitor is to have a capacitance of 1.25×10^{-9} F and must be able to withstand a maximum potential difference of 5500 V. What is the minimum area the plates of the capacitor may have?

24.38 •• BIO **Potential in Human Cells.** Some cell walls in the human body have a layer of negative charge on the inside surface and a layer of positive charge of equal magnitude on the outside surface. Suppose that the charge density on either surface is $\pm 0.50 \times 10^{-3}$ C/m^2, the cell wall is 5.0 nm thick, and the cell-wall material is air. (a) Find the magnitude of $\vec{E}$ in the wall between the two layers of charge. (b) Find the potential difference between the inside and the outside of the cell. Which is at the higher potential? (c) A typical cell in the human body has a volume of 10^{-16} m^3. Estimate the total electric-field energy stored in the wall of a cell of this size. (*Hint:* Assume that the cell is spherical, and calculate the volume of the cell wall.) (d) In reality, the cell wall is made up, not of air, but of tissue with a dielectric constant of 5.4. Repeat parts (a) and (b) in this case.

24.39 • A constant potential difference of 12 V is maintained between the terminals of a 0.25 μF, parallel-plate, air capacitor. (a) A sheet of Mylar is inserted between the plates of the capacitor, completely filling the space between the plates. When this is done, how much additional charge flows onto the positive plate of the capacitor (see Table 24.1)? (b) What is the total induced charge on either face of the Mylar sheet? (c) What effect does the Mylar sheet have on the electric field between the plates? Explain how you can reconcile this with the increase in charge on the plates, which acts to *increase* the electric field.

24.40 •• Polystyrene has dielectric constant 2.6 and dielectric strength 2.0×10^7 V/m. A piece of polystyrene is used as a dielectric in a parallel-plate capacitor, filling the volume between the plates. (a) When the electric field between the plates is 80% of the dielectric strength, what is the energy density of the stored energy? (b) When the capacitor is connected to a battery with voltage 500.0 V, the electric field between the plates is 80% of the dielectric strength. What is the area of each plate if the capacitor stores 0.200 mJ of energy under these conditions?

24.41 • When a 360 nF air capacitor $(1\text{ nF} = 10^{-9}\text{ F})$ is connected to a power supply, the energy stored in the capacitor is 1.85×10^{-5} J. While the capacitor is kept connected to the power supply, a slab of dielectric is inserted that completely fills the space between the plates. This increases the stored energy by 2.32×10^{-5} J. (a) What is the potential difference between the capacitor plates? (b) What is the dielectric constant of the slab?

24.42 • A parallel-plate capacitor has capacitance $C = 12.5$ pF when the volume between the plates is filled with air. The plates are circular, with radius 3.00 cm. The capacitor is connected to a battery, and a charge of magnitude 25.0 pC goes onto each plate. With the capacitor still connected to the battery, a slab of dielectric is inserted between the plates, completely filling the space between the plates. After the dielectric has been inserted, the charge on each plate has magnitude 45.0 pC. (a) What is the dielectric constant K of the dielectric? (b) What is the potential difference between the plates before and after the dielectric has been inserted? (c) What is the electric field at a point midway between the plates before and after the dielectric has been inserted?

Section 24.6 Gauss's Law in Dielectrics

24.43 • A parallel-plate capacitor has the volume between its plates filled with plastic with dielectric constant K. The magnitude of the charge on each plate is Q. Each plate has area A, and the distance between the plates is d. (a) Use Gauss's law as stated in Eq. (24.23) to calculate the magnitude of the electric field in the dielectric. (b) Use the electric field determined in part (a) to calculate the potential difference between the two plates. (c) Use the result of part (b) to determine the capacitance of the capacitor. Compare your result to Eq. (24.12).

24.44 • A parallel-plate capacitor has plates with area 0.0225 m^2 separated by 1.00 mm of Teflon. (a) Calculate the charge on the plates when they are charged to a potential difference of 12.0 V. (b) Use Gauss's law [Eq. (24.23)] to calculate the electric field inside the Teflon. (c) Use Gauss's law to calculate the electric field if the voltage source is disconnected and the Teflon is removed.

PROBLEMS

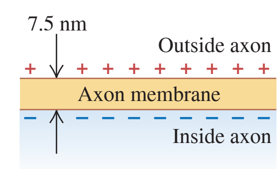
24.45 • Electronic flash units for cameras contain a capacitor for storing the energy used to produce the flash. In one such unit, the flash lasts for $\frac{1}{675}$ s with an average light power output of 2.70×10^5 W. (a) If the conversion of electrical energy to light is 95% efficient (the rest of the energy goes to thermal energy), how much energy must be stored in the capacitor for one flash? (b) The capacitor has a potential difference between its plates of 125 V when the stored energy equals the value calculated in part (a). What is the capacitance?

24.46 • A parallel-plate air capacitor is made by using two plates 12 cm square, spaced 3.7 mm apart. It is connected to a 12 V battery. (a) What is the capacitance? (b) What is the charge on each plate? (c) What is the electric field between the plates? (d) What is the energy stored in the capacitor? (e) If the battery is disconnected and then the plates are pulled apart to a separation of 7.4 mm, what are the answers to parts (a)–(d)?

24.47 ••• In one type of computer keyboard, each key holds a small metal plate that serves as one plate of a parallel-plate, air-filled capacitor. When the key is depressed, the plate separation decreases and the capacitance increases. Electronic circuitry detects the change in capacitance and thus detects that the key has been pressed. In one particular keyboard, the area of each metal plate is 42.0 mm^2, and the separation between the plates is 0.700 mm before the key is depressed. (a) Calculate the capacitance before the key is depressed. (b) If the circuitry can detect a change in capacitance of 0.250 pF, how far must the key be depressed before the circuitry detects its depression?

24.48 ••• BIO **Cell Membranes.** Cell membranes (the walled enclosure around a cell) are typically about 7.5 nm thick. They are partially permeable to allow charged material to pass in and out, as needed. Equal but opposite charge densities build up on the inside and outside faces of such a membrane, and these charges prevent additional charges from passing through the cell wall. We can model a cell membrane as a parallel-plate capacitor, with the membrane itself containing proteins embedded in an organic material to give the membrane a dielectric constant of about 10. (See **Fig. P24.48**.) (a) What is the capacitance per square centimeter of such a cell wall? (b) In its normal resting state, a cell has a potential difference of 85 mV across its membrane. What is the electric field inside this membrane?

Figure P24.48

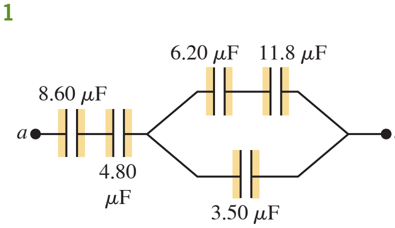


24.49 •• A 20.0 μF capacitor is charged to a potential difference of 800 V. The terminals of the charged capacitor are then connected to those of an uncharged 10.0 μF capacitor. Compute (a) the original charge of the system, (b) the final potential difference across each capacitor, (c) the final energy of the system, and (d) the decrease in energy when the capacitors are connected.

24.50 •• When lightning strikes a car, the metallic outer shell, which is insulated from the ground by its rubber tires, attains a high voltage. We can estimate how much charge is deposited by roughly modeling the car as a spherical capacitor with the outer radius taken to infinity. (a) Determine the capacitance of a sphere in terms of its radius, either by considering the potential on a sphere relative to infinity as a function of its charge, or by considering a spherical capacitor as the outer shell becomes very large. (These methods provide the same result.) (b) Estimate the radius of a sphere that corresponds to the size of a car. (c) Determine the corresponding capacitance. (d) A typical lighting strike could deliver a 100 MV potential to a car. Then what net charge would be deposited on the car?

24.51 • For the capacitor network shown in **Fig. P24.51**, the potential difference across ab is 12.0 V. Find (a) the total energy stored in this network and (b) the energy stored in the 4.80 μF capacitor.

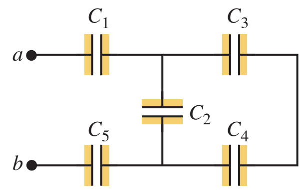
Figure P24.51



24.52 •• In Fig. E24.17, $C_1 = 6.00\ \mu$F, $C_2 = 3.00\ \mu$F, $C_3 = 4.00\ \mu$F, and $C_4 = 8.00\ \mu$F. The capacitor network is connected to an applied potential difference V_{ab}. After the charges on the capacitors have reached their final values, the voltage across C_3 is 40.0 V. What are (a) the voltages across C_1 and C_2, (b) the voltage across C_4, and (c) the voltage V_{ab} applied to the network?

24.53 • In **Fig. P24.53**, $C_1 = C_5 = 8.4\ \mu F$ and $C_2 = C_3 = C_4 = 4.2\ \mu F$. The applied potential is $V_{ab} = 220$ V. (a) What is the equivalent capacitance of the network between points a and b? (b) Calculate the charge on each capacitor and the potential difference across each capacitor.

Figure **P24.53**



24.54 •• **CP** After combing your hair on a dry day, some of your hair stands up, forced away from your head by electrostatic repulsion. (a) Estimate the length L of your hair. (b) Using the average linear mass density of hair, which is 65 μg/cm, estimate the mass m of one of your hairs. (c) Estimate the number N of hairs that stand after combing. (d) Assume that the comb has taken away a charge $-2Q$, and that your hair has therefore gained an amount of charge $2Q$. Assume further that half of this charge resides next to your head and the other half is distributed at the ends of the N strands that stand up. Assume the electrostatic force that lifted a hair was twice its weight. Show that this leads to

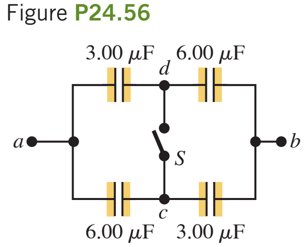
$$2mg = \frac{1}{4\pi\epsilon_0}\frac{Q^2/N}{L^2}$$

Use this equation to estimate the charge Q that resides on your head. (e) If your head were a sphere with radius R, it would have a capacitance of $4\pi\epsilon_0 R$. Estimate the radius of your head; then use your estimate to determine your head's capacitance. (f) Use your result to estimate the potential attained due to combing. (The surprising result illustrates an interesting feature of static electricity.)

24.55 •• In Fig. E24.20, $C_1 = 3.00\ \mu F$ and $V_{ab} = 150$ V. The charge on capacitor C_1 is 150 μC and the charge on C_3 is 450 μC. What are the values of the capacitances of C_2 and C_3?

24.56 • The capacitors in **Fig. P24.56** are initially uncharged and are connected, as in the diagram, with switch S open. The applied potential difference is $V_{ab} = +210$ V. (a) What is the potential difference V_{cd}? (b) What is the potential difference across each capacitor after switch S is closed? (c) How much charge flowed through the switch when it was closed?

Figure **P24.56**

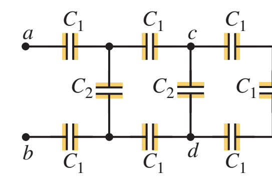


24.57 •• Three capacitors having capacitances of 8.4, 8.4, and 4.2 μF are connected in series across a 36 V potential difference. (a) What is the charge on the 4.2 μF capacitor? (b) What is the total energy stored in all three capacitors? (c) The capacitors are disconnected from the potential difference without allowing them to discharge. They are then reconnected in parallel with each other, with the positively charged plates connected together. What is the voltage across each capacitor in the parallel combination? (d) What is the total energy now stored in the capacitors?

24.58 • **Capacitance of a Thundercloud.** The charge center of a thundercloud, drifting 3.0 km above the earth's surface, contains 20 C of negative charge. Assuming the charge center has a radius of 1.0 km, and modeling the charge center and the earth's surface as parallel plates, calculate: (a) the capacitance of the system; (b) the potential difference between charge center and ground; (c) the average strength of the electric field between cloud and ground; (d) the electrical energy stored in the system.

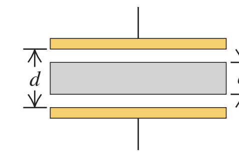
24.59 •• In **Fig. P24.59**, each capacitance C_1 is 6.9 μF, and each capacitance C_2 is 4.6 μF. (a) Compute the equivalent capacitance of the network between points a and b. (b) Compute the charge on each of the three capacitors nearest a and b when $V_{ab} = 420$ V. (c) With 420 V across a and b, compute V_{cd}.

Figure **P24.59**



24.60 •• An air capacitor is made by using two flat plates, each with area A, separated by a distance d. Then a metal slab having thickness a (less than d) and the same shape and size as the plates is inserted between them, parallel to the plates and not touching either plate (**Fig. P24.60**). (a) What is the capacitance of this arrangement? (b) Express the capacitance as a multiple of the capacitance C_0 when the metal slab is not present. (c) Discuss what happens to the capacitance in the limits $a \rightarrow 0$ and $a \rightarrow d$.

Figure **P24.60**



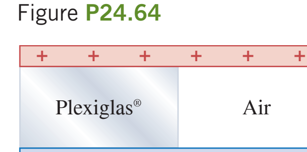
24.61 •• A potential difference $V_{ab} = 48.0$ V is applied across the capacitor network of Fig. E24.17. If $C_1 = C_2 = 4.00\ \mu F$ and $C_4 = 8.00\ \mu F$, what must the capacitance C_3 be if the network is to store 2.90×10^{-3} J of electrical energy?

24.62 • **CALC** The inner cylinder of a long, cylindrical capacitor has radius r_a and linear charge density $+\lambda$. It is surrounded by a coaxial cylindrical conducting shell with inner radius r_b and linear charge density $-\lambda$ (see Fig. 24.6). (a) What is the energy density in the region between the conductors at a distance r from the axis? (b) Integrate the energy density calculated in part (a) over the volume between the conductors in a length L of the capacitor to obtain the total electric-field energy per unit length. (c) Use Eq. (24.9) and the capacitance per unit length calculated in Example 24.4 (Section 24.1) to calculate U/L. Does your result agree with that obtained in part (b)?

24.63 ••• **CP** Dielectric elastomers are used to create voltage-dependent capacitors. These elastomers contract when subject to the stresses induced by Coulomb forces between capacitor plates. Consider a pair of parallel plates with area A separated by distance d_0 when no voltage is applied. A material with dielectric constant K and Young's modulus Y fills the space between the plates. (a) What compressive force is applied to the dielectric when the plates have opposite charges Q and $-Q$? (b) The distance between the plates is $d = d_0 - sQ^2$, where s is a squeezing coefficient. Determine s in terms of the parameters specified. (c) A capacitor of this sort has plate area 1.00 cm^2 and noncharged separation distance $d_0 = 0.400$ mm. The plates are separated by a silicone dielectric elastomer with dielectric constant $K = 3.00$ and Young's modulus 0.0100 GPa. What is the squeezing coefficient in this case? (d) If each plate has charge of magnitude 0.700 μC, what is the applied voltage? (e) What new voltage would double the amount of charge on the plates?

24.64 •• A parallel-plate capacitor is made from two plates 12.0 cm on each side and 4.50 mm apart. Half of the space between these plates contains only air, but the other half is filled with Plexiglas® of dielectric constant 3.40 (**Fig. P24.64**). An 18.0 V battery is connected across the plates. (a) What is the capacitance of this combination? (*Hint:* Can you think of this capacitor as equivalent to two capacitors in parallel?) (b) How much energy is stored in the capacitor? (c) If we remove the Plexiglas but change nothing else, how much energy will be stored in the capacitor?

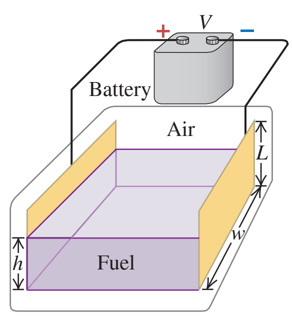
Figure **P24.64**



24.65 ••• CP CALC A metallic circular plate with radius r is fixed to a tabletop. An identical circular plate supported from above by a cable is fixed in place a distance d above the first plate. Assume that d is much smaller than r. The two plates are attached by wires to a battery that supplies voltage V. (a) What is the tension in the cable? Neglect the weight of the plate. (b) The upper plate is slowly raised to a new height $2d$. Determine the work done by the cable by integrating $\int_d^{2d} F(z)dz$, where $F(z)$ is the cable tension when the plates are separated by a distance z. (c) Compute the energy stored in the electric field before the top plate was raised. (d) Compute the energy stored in the electric field after the top plate was raised. (e) Is the work done by the cable equal to the change in the stored electrical energy? If not, why not?

24.66 •• A fuel gauge uses a capacitor to determine the height of the fuel in a tank. The effective dielectric constant K_{eff} changes from a value of 1 when the tank is empty to a value of K, the dielectric constant of the fuel, when the tank is full. The appropriate electronic circuitry can determine the effective dielectric constant of the combined air and fuel between the capacitor plates. Each of the two rectangular plates has a width w and a length L (**Fig. P24.66**). The height of the fuel between the plates is h. You can ignore any fringing effects. (a) Derive an expression for K_{eff} as a function of h. (b) What is the effective dielectric constant for a tank $\frac{1}{4}$ full, $\frac{1}{2}$ full, and $\frac{3}{4}$ full if the fuel is gasoline ($K = 1.95$)? (c) Repeat part (b) for methanol ($K = 33.0$). (d) For which fuel is this fuel gauge more practical?

Figure **P24.66**

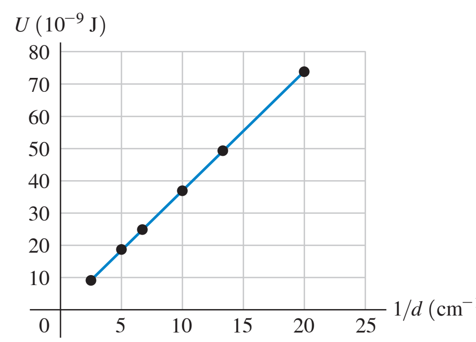


24.67 •• DATA Your electronics company has several identical capacitors with capacitance C_1 and several others with capacitance C_2. You must determine the values of C_1 and C_2 but don't have access to C_1 and C_2 individually. Instead, you have a network with C_1 and C_2 connected in series and a network with C_1 and C_2 connected in parallel. You have a 200.0 V battery and instrumentation that measures the total energy supplied by the battery when it is connected to the network. When the parallel combination is connected to the battery, 0.180 J of energy is stored in the network. When the series combination is connected, 0.0400 J of energy is stored. You are told that C_1 is greater than C_2. (a) Calculate C_1 and C_2. (b) For the series combination, does C_1 or C_2 store more charge, or are the values equal? Does C_1 or C_2 store more energy, or are the values equal? (c) Repeat part (b) for the parallel combination.

24.68 •• DATA You are designing capacitors for various applications. For one application, you want the maximum possible stored energy. For another, you want the maximum stored charge. For a third application, you want the capacitor to withstand a large applied voltage without dielectric breakdown. You start with an air-filled parallel-plate capacitor that has $C_0 = 6.00$ pF and a plate separation of 2.50 mm. You then consider the use of each of the dielectric materials listed in Table 24.2. In each application, the dielectric will fill the volume between the plates, and the electric field between the plates will be 50% of the dielectric strength given in the table. (a) For each of the five materials given in the table, calculate the energy stored in the capacitor. Which dielectric allows the maximum stored energy? (b) For each material, what is the charge Q stored on each plate of the capacitor? (c) For each material, what is the voltage applied across the capacitor? (d) Is one dielectric material in the table your best choice for all three applications?

24.69 •• DATA You are conducting experiments with an air-filled parallel-plate capacitor. You connect the capacitor to a battery with voltage 24.0 V. Initially the separation d between the plates is 0.0500 cm. In one experiment, you leave the battery connected to the capacitor, increase the separation between the plates, and measure the energy stored in the capacitor for each value of d. In a second experiment, you make the same measurements but disconnect the battery before you change the plate separation. One set of your data is given in **Fig. P24.69**, where you have plotted the stored energy U versus $1/d$. (a) For which experiment does this data set apply: the first (battery remains connected) or the second (battery disconnected before d is changed)? Explain. (b) Use the data plotted in Fig. P24.69 to calculate the area A of each plate. (c) For which case, the battery connected or the battery disconnected, is there more energy stored in the capacitor when $d = 0.400$ cm? Explain.

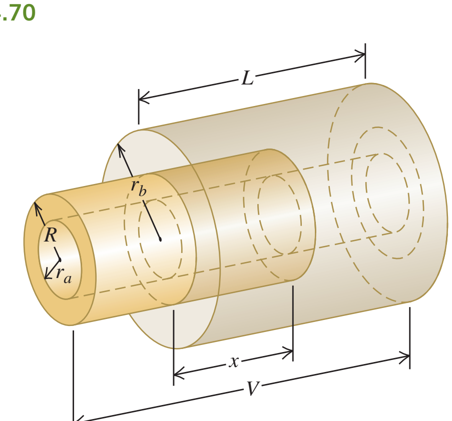
Figure **P24.69**



CHALLENGE PROBLEMS

24.70 ••• A thin-walled, hollow, conducting cylinder with radius r_b is concentric with a solid conducting cylinder with radius $r_a < r_b$. Each has length L. The two cylinders are attached by conducting wires to the terminals of a battery that supplies potential V. A solid cylindrical shell, with inner radius r_a and outer radius $R < r_b$, made of a material with dielectric constant K, slides between the conducting cylinders, as shown in **Fig. P24.70**. By changing the insertion distance x, we can alter the capacitance seen by the battery and therefore alter the amount of charge stored in this device. (a) Determine the capacitance as a function of x. (b) If $L = 10.0$ cm, $r_a = 1.00$ cm, $r_b = 4.00$ cm, $R = 3.00$ cm, and $K = 3.21$, what is the capacitance when $x = 0$? (c) What is the capacitance when $x = L$? (d) What value of x results in 6.00 nC of charge on the positively charged cylinder plate when $V = 1.00$ kV?

Figure **P24.70**

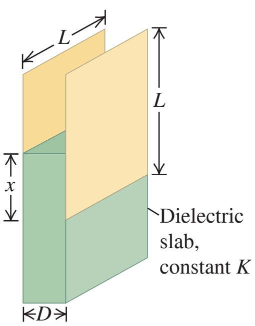


24.71 ••• **CALC** Two conducting plates with area A are separated by distance d. Between the plates is a material with a dielectric constant that varies linearly from a value of unity next to one plate to K next to the other plate. (a) What is the capacitance of this device? [*Hint:* We can envision this as a continuum of capacitors with differential plate separation connected in series. The reciprocal of the capacitance of a differential slice is then $dx/(K(x)\epsilon_0 A)$, where $K(x)$ is the dielectric constant specific to that locale.] (b) Show that this result matches the expected result when $K \rightarrow 1$.

24.72 ••• Two square conducting plates with sides of length L are separated by a distance D. A dielectric slab with constant K with dimensions $L \times L \times D$ is inserted a distance x into the space between the plates, as shown in **Fig. P24.72**. (a) Find the capacitance C of this system. (b) Suppose that the capacitor is connected to a battery that maintains a constant potential difference V between the plates. If the dielectric slab is inserted an additional distance dx into the space between the plates, show that the change in stored energy is

$$dU = +\frac{(K-1)\epsilon_0 V^2 L}{2D}dx$$

Figure **P24.72**



(c) Suppose that before the slab is moved by dx, the plates are disconnected from the battery, so that the charges on the plates remain constant. Determine the magnitude of the charge on each plate, and then show that when the slab is moved dx farther into the space between the plates, the stored energy changes by an amount that is the *negative* of the expression for dU given in part (b). (d) If F is the force exerted on the slab by the charges on the plates, then dU should equal the work done *against* this force to move the slab a distance dx. Thus $dU = -F\,dx$. Show that applying this expression to the result of part (b) suggests that the electric force on the slab pushes it *out* of the capacitor, while the result of part (c) suggests that the force pulls the slab *into* the capacitor. (e) Figure 24.16 shows that the force in fact pulls the slab into the capacitor. Explain why the result of part (b) gives an incorrect answer for the direction of this force, and calculate the magnitude of the force. (This method does not require knowledge of the nature of the fringing field.)

MCAT-STYLE PASSAGE PROBLEMS

BIO The Electric Egg. Upon fertilization, the eggs of many species undergo a rapid change in potential difference across their outer membrane. This change affects the physiological development of the eggs. The potential difference across the membrane is called the *membrane potential,* V_m, which is the potential inside the membrane minus the potential outside it. The membrane potential arises when enzymes use the energy available in ATP to expel three sodium ions (Na^+) actively and accumulate two potassium ions (K^+) inside the membrane—making the interior less positively charged than the exterior. For a sea urchin egg, V_m is about -70 mV; that is, the potential inside is 70 mV less than that outside. The egg membrane behaves as a capacitor with a capacitance of about 1 $\mu F/cm^2$. The membrane of the unfertilized egg is *selectively permeable* to K^+; that is, K^+ can readily pass through certain channels in the membrane, but other ions cannot. When a sea urchin egg is fertilized, Na^+ channels in the membrane open, Na^+ enters the egg, and V_m rapidly increases to $+30$ mV, where it remains for several minutes. The concentration of Na^+ is about 30 mmol/L in the egg's interior but 450 mmol/L in the surrounding seawater. The K^+ concentration is about 200 mmol/L inside but 10 mmol/L outside. A useful constant that connects electrical and chemical units is the *Faraday number,* which has a value of approximately 10^5 C/mol; that is, Avogadro's number (a mole) of monovalent ions, such as Na^+ or K^+, carries a charge of 10^5 C.

24.73 How many moles of Na^+ must move per unit area of membrane to change V_m from -70 mV to $+30$ mV, if we assume that the membrane behaves purely as a capacitor? (a) 10^{-4} mol/cm^2; (b) 10^{-9} mol/cm^2; (c) 10^{-12} mol/cm^2; (d) 10^{-14} mol/cm^2.

24.74 Suppose that the egg has a diameter of 200 μm. What fractional change in the internal Na^+ concentration results from the fertilization-induced change in V_m? Assume that Na^+ ions are distributed throughout the cell volume. The concentration increases by (a) 1 part in 10^4; (b) 1 part in 10^5; (c) 1 part in 10^6; (d) 1 part in 10^7.

24.75 Suppose that the change in V_m was caused by the entry of Ca^{2+} instead of Na^+. How many Ca^{2+} ions would have to enter the cell per unit membrane to produce the change? (a) Half as many as for Na^+; (b) the same as for Na^+; (c) twice as many as for Na^+; (d) cannot say without knowing the inside and outside concentrations of Ca^{2+}.

24.76 What is the minimum amount of work that must be done by the cell to restore V_m to -70 mV? (a) 3 mJ; (b) 3 μJ; (c) 3 nJ; (d) 3 pJ.

ANSWERS

Chapter Opening Question ?

(iv) Equation (24.9) shows that the energy stored in a capacitor with capacitance C and charge Q is $U = Q^2/2C$. If Q is doubled, the stored energy increases by a factor of $2^2 = 4$. Note that if the value of Q is too great, the electric-field magnitude inside the capacitor will exceed the dielectric strength of the material between the plates and dielectric breakdown will occur (see Section 24.4). This puts a practical limit on the amount of energy that can be stored.

Key Example VARIATION Problems

VP24.4.1 (a) 6.58×10^{-9} F (b) 5.85 kV
VP24.4.2 (a) 4.80 mm (b) $Q = 2.49 \times 10^{-5}$ C, $E = 6.25 \times 10^5$ V/m
VP24.4.3 (a) 8.33 cm (b) 4.44×10^{-8} C (c) $\sigma_{inner} = +5.09 \times 10^{-7}$ C/m^2, $\sigma_{outer} = -4.36 \times 10^{-7}$ C/m^2
VP24.4.4 (a) 2.24 (b) 125 V; inner conductor
VP24.9.1 (a) 0.625 μF (b) 8.50 μF (c) 2.06 μF (d) 5.71 μF
VP24.9.2 (a) $U_1 = 0.0675$ J, $U_2 = 0.0169$ J; C_1 (b) $U_1 = 0.0117$ J, $U_2 = 0.0469$ J; C_2
VP24.9.3 (a) series 0.0115 J, parallel 0.0578 J; parallel (c) series 0.0733 J, parallel 0.0145 J; series
VP24.9.4 (a) 5.10×10^{-6} C/m^2 (b) 5.76×10^5 V/m (c) 1.47 J/m^3 (d) 5.04×10^{-3} J
VP24.11.1 (a) 1.36 nF (b) 5.43×10^{-6} C (c) 1.60 (d) 2.17 nF (e) 2.04×10^{-6} C
VP24.11.2 (a) 2.36 nF (b) 8.26×10^{-6} C (c) 5.90 nF (d) 2.07×10^{-5} C (e) 1.24×10^{-5} C
VP24.11.3 (a) 17.6 J (b) 7.56 J
VP24.11.4 (a) 6.30 J (b) 18.0 J

Bridging Problem

(a) 0
(b) $Q^2/32\pi^2\epsilon_0 r^4$
(c) $Q^2/8\pi\epsilon_0 R$
(d) $Q^2/8\pi\epsilon_0 R$
(e) $C = 4\pi\epsilon_0 R$

? In a flashlight, how does the amount of current that flows out of the bulb compare to the amount that flows into the bulb? (i) Current out is less than current in; (ii) current out is greater than current in; (iii) current out equals current in; (iv) the answer depends on the brightness of the bulb.

25 Current, Resistance, and Electromotive Force

LEARNING OUTCOMES

In this chapter, you'll learn...

25.1 The meaning of electric current, and how charges move in a conductor.
25.2 What is meant by the resistivity and conductivity of a substance.
25.3 How to calculate the resistance of a conductor from its dimensions and its resistivity.
25.4 How an electromotive force (emf) makes it possible for current to flow in a circuit.
25.5 How to do calculations involving energy and power in circuits.
25.6 How to use a simple model to understand the flow of current in metals.

You'll need to review...

17.7 Thermal conductivity.
23.2 Voltmeters, electric field, and electric potential.
24.4 Dielectric breakdown in insulators.

In the past four chapters we studied the interactions of electric charges *at rest;* now we're ready to study charges *in motion*. An *electric current* consists of charges in motion from one region to another. If the charges follow a conducting path that forms a closed loop, the path is called an *electric circuit*.

Fundamentally, electric circuits are a means for conveying *energy* from one place to another. As charged particles move within a circuit, electric potential energy is transferred from a source (such as a battery or generator) to a device in which that energy is either stored or converted to another form: into sound in a stereo system or into heat and light in a toaster or light bulb. Electric circuits are useful because they allow energy to be transported without any moving parts (other than the moving charged particles themselves). They are at the heart of computers, television transmitters and receivers, and household and industrial power distribution systems. Your nervous system is a specialized electric circuit that carries vital signals from one part of your body to another.

In Chapter 26 we'll see how to analyze electric circuits and examine some practical applications of circuits. To prepare you for that, in this chapter we'll examine the basic properties of electric currents. We'll begin by describing the nature of electric conductors and considering how they are affected by temperature. We'll learn why a short, fat, cold copper wire is a better conductor than a long, skinny, hot steel wire. We'll study the properties of batteries and see how they cause current and energy transfer in a circuit. In this analysis we'll use the concepts of current, potential difference (or voltage), resistance, and electromotive force. Finally, we'll look at electric current in a material from a microscopic viewpoint.

25.1 CURRENT

A **current** is any motion of charge from one region to another. In this section we'll discuss currents in conducting materials. The vast majority of technological applications of charges in motion involve currents of this kind.

In electrostatic situations (discussed in Chapters 21 through 24) the electric field is zero everywhere within the conductor, and there is *no* current. However, this does not mean that all charges within the conductor are at rest. In an ordinary metal such as copper or aluminum, some of the electrons are free to move within the conducting material. These free electrons move randomly in all directions, somewhat like the molecules of a gas but with much greater speeds, of the order of 10^6 m/s. The electrons nonetheless do not escape from the conducting material, because they are attracted to the positive ions of the material. The motion of the electrons is random, so there is no *net* flow of charge in any direction and hence no current.

Now consider what happens if a constant, steady electric field $\vec{E}$ is established inside a conductor. (We'll see later how this can be done.) A charged particle (such as a free electron) inside the conducting material is then subjected to a steady force $\vec{F} = q\vec{E}$. If the charged particle were moving in *vacuum,* this steady force would cause a steady acceleration in the direction of $\vec{F}$, and after a time the charged particle would be moving in that direction at high speed. But a charged particle moving in a *conductor* undergoes frequent collisions with the massive, nearly stationary ions of the material. In each such collision the particle's direction of motion undergoes a random change. The net effect of the electric field $\vec{E}$ is that in addition to the random motion of the charged particles within the conductor, there is also a very slow net motion or *drift* of the moving charged particles as a group in the direction of the electric force $\vec{F} = q\vec{E}$ (**Fig. 25.1**). This motion is described in terms of the **drift velocity** $\vec{v}_d$ of the particles. As a result, there is a net current in the conductor.

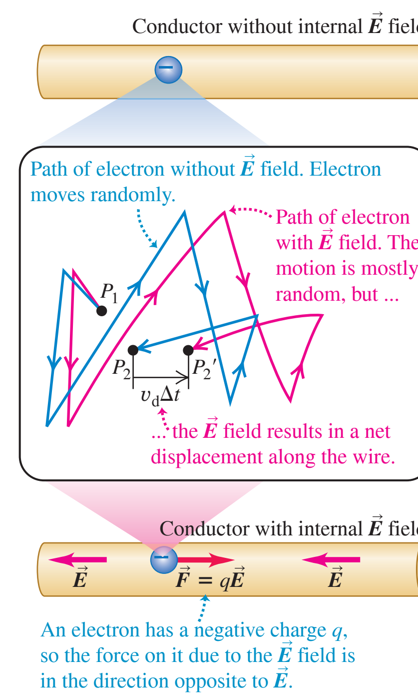


Figure **25.1** If there is no electric field inside a conductor, an electron moves randomly from point P_1 to point P_2 in a time Δt. If an electric field $\vec{E}$ is present, the electric force $\vec{F} = q\vec{E}$ imposes a small drift (greatly exaggerated here) that takes the electron to point P_2', a distance $v_d \Delta t$ from P_2 in the direction of the force.

While the random motion of the electrons has a very fast average speed of about 10^6 m/s, the drift speed is very slow, often on the order of 10^{-4} m/s. Given that the electrons move so slowly, you may wonder why the light comes on immediately when you turn on the switch of a flashlight. The reason is that the electric field is set up in the wire with a speed approaching the speed of light, and electrons start to move all along the wire at very nearly the same time. The time that it takes any individual electron to get from the switch to the light bulb isn't really relevant. A good analogy is a group of soldiers standing at attention when the sergeant orders them to start marching; the order reaches the soldiers' ears at the speed of sound, which is much faster than their marching speed, so all the soldiers start to march essentially in unison.

The Direction of Current Flow

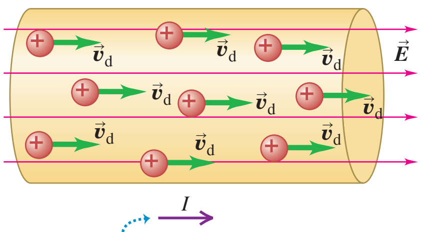
The drift of moving charges through a conductor can be interpreted in terms of work and energy. The electric field $\vec{E}$ does work on the moving charges. The resulting kinetic energy is transferred to the material of the conductor by means of collisions with the ions, which vibrate about their equilibrium positions in the crystalline structure of the conductor. This energy transfer increases the average vibrational energy of the ions and therefore the temperature of the material. Thus much of the work done by the electric field goes into heating the conductor, *not* into making the moving charges move ever faster and faster. This heating is sometimes useful, as in an electric toaster, but in many situations is simply an unavoidable by-product of current flow.

In different current-carrying materials, the charges of the moving particles may be positive or negative. In metals the moving charges are always (negative) electrons, while in an ionized gas (plasma) or an ionic solution the moving charges may include both electrons and positively charged ions. In a semiconductor material such as germanium or silicon, conduction is partly by electrons and partly by motion of *vacancies,* also known as *holes;* these are sites of missing electrons and act like positive charges.

Figure 25.2 (next page) shows segments of two different current-carrying materials. In Fig. 25.2a the moving charges are positive, the electric force is in the same direction as

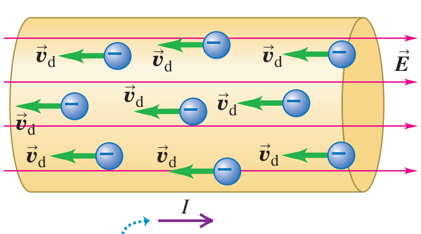
Figure **25.2** The same current is produced by (a) positive charges moving in the direction of the electric field $\vec{E}$ or (b) the same number of negative charges moving at the same speed in the direction opposite to $\vec{E}$.

(a)



A **conventional current** is treated as a flow of positive charges, regardless of whether the free charges in the conductor are positive, negative, or both.

(b)



In a metallic conductor, the moving charges are electrons — but the *current* still points in the direction positive charges would flow.

$\vec{E}$, and the drift velocity $\vec{v}_d$ is from left to right. In Fig. 25.2b the charges are negative, the electric force is opposite to $\vec{E}$, and the drift velocity $\vec{v}_d$ is from right to left. In both cases there is a net flow of positive charge from left to right, and positive charges end up to the right of negative ones. We *define* the current, denoted by I, to be in the direction in which there is a flow of *positive* charge. Thus we describe currents as though they consisted entirely of positive charge flow, even in cases in which we know that the actual current is due to electrons. Hence the current is to the right in both Figs. 25.2a and 25.2b. This choice or convention for the direction of current flow is called **conventional current.** While the direction of the conventional current is *not* necessarily the same as the direction in which charged particles are actually moving, we'll find that the sign of the moving charges is of little importance in analyzing electric circuits.

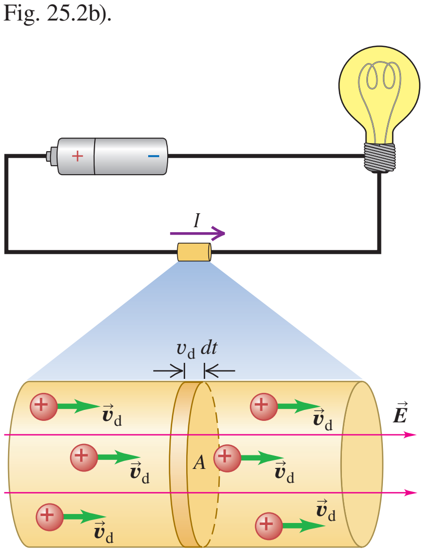
Figure 25.3 shows a segment of a conductor in which a current is flowing. We consider the moving charges to be *positive,* so they are moving in the same direction as the current. We define the current through the cross-sectional area A to be *the net charge flowing through the area per unit time*. Thus, if a net charge dQ flows through an area in a time dt, the current I through the area is

$$I = \frac{dQ}{dt} \qquad \text{(definition of current)} \qquad (25.1)$$

CAUTION **Current is not a vector** Although we refer to the *direction* of a current, current as defined by Eq. (25.1) is *not* a vector quantity. In a current-carrying wire, the current is always along the length of the wire, regardless of whether the wire is straight or curved. No single vector could describe motion along a curved path. We'll usually describe the direction of current either in words (as in "the current flows clockwise around the circuit") or by choosing a current to be positive if it flows in one direction along a conductor and negative if it flows in the other direction.

The SI unit of current is the **ampere;** one ampere is defined to be *one coulomb per second* $(1\ \text{A} = 1\ \text{C/s})$. This unit is named in honor of the French scientist André Marie Ampère (1775–1836). When an ordinary flashlight (D-cell size) is turned on, the current in the flashlight is about 0.5 to 1 A; the current in the wires of a car engine's starter motor is around 200 A. Currents in radio and television circuits are usually expressed in *milliamperes* $(1\ \text{mA} = 10^{-3}\ \text{A})$ or *microamperes* $(1\ \mu\text{A} = 10^{-6}\ \text{A})$, and currents in computer circuits are expressed in *nanoamperes* $(1\ \text{nA} = 10^{-9}\ \text{A})$ or *picoamperes* $(1\ \text{pA} = 10^{-12}\ \text{A})$.

Figure **25.3** The current I is the time rate of charge transfer through the cross-sectional area A. The random component of each moving charged particle's motion averages to zero, and the current is in the same direction as $\vec{E}$ whether the moving charges are positive (as shown here) or negative (see Fig. 25.2b).



Current $I = \frac{dQ}{dt}$

Current, Drift Velocity, and Current Density

We can express current in terms of the drift velocity of the moving charges. Let's reconsider the situation of Fig. 25.3 of a conductor with cross-sectional area A and an electric field $\vec{E}$ directed from left to right. To begin with, we'll assume that the free charges in the conductor are positive; then the drift velocity is in the same direction as the field.

Suppose there are n moving charged particles per unit volume. We call n the **concentration** of particles; its SI unit is m^{-3}. Assume that all the particles move with the same drift velocity with magnitude v_d. In a time interval dt, each particle moves a distance $v_d\,dt$. The particles that flow out of the right end of the shaded cylinder with length $v_d\,dt$ during dt are the particles that were within this cylinder at the beginning of the interval dt. The volume of the cylinder is $Av_d\,dt$, and the number of particles within it is $nAv_d\,dt$. If each particle has a charge q, the charge dQ that flows out of the end of the cylinder during time dt is

$$dQ = q(nAv_d\,dt) = nqv_d A\,dt$$

and the current is

$$I = \frac{dQ}{dt} = nqv_d A$$

The current *per unit cross-sectional area* is called the **current density** J:

$$J = \frac{I}{A} = nqv_{\mathrm{d}}$$

The units of current density are amperes per square meter $(\mathrm{A/m^2})$.

If the moving charges are negative rather than positive, as in Fig. 25.2b, the drift velocity is opposite to $\vec{E}$. But the *current* is still in the same direction as $\vec{E}$ at each point in the conductor. Hence current I and current density J don't depend on the sign of the charge, and so we replace the charge q by its absolute value $|q|$:

Rate at which charge flows through area

Current through an area

$$I = \frac{dQ}{dt} = n|q|v_{\mathrm{d}}A \qquad (25.2)$$

Drift speed

Cross-sectional area

Concentration of moving charged particles

Charge per particle

$$J = \frac{I}{A} = n|q|v_{\mathrm{d}} \qquad \text{(current density)} \qquad (25.3)$$

The current in a conductor is the product of the concentration of moving charged particles, the magnitude of charge of each such particle, the magnitude of the drift velocity, and the cross-sectional area of the conductor.

We can also define a *vector* current density $\vec{J}$ that includes the direction of the drift velocity:

Vector current density

$$\vec{\boldsymbol{J}} = nq\vec{\boldsymbol{v}}_{\mathrm{d}} \qquad (25.4)$$

Drift velocity

Concentration of moving charged particles

Charge per particle

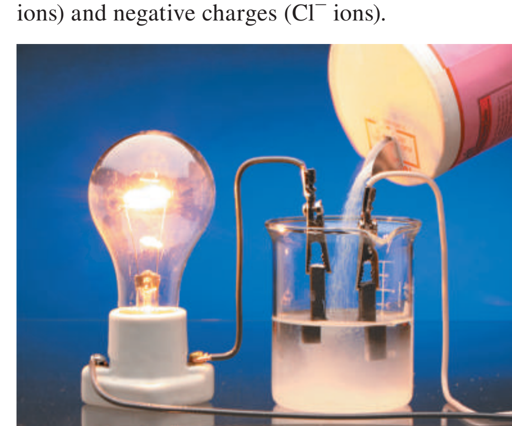
There are *no* absolute value signs in Eq. (25.4). If q is positive, $\vec{v}_{\mathrm{d}}$ is in the same direction as $\vec{E}$; if q is negative, $\vec{v}_{\mathrm{d}}$ is opposite to $\vec{E}$. In either case, $\vec{J}$ is in the same direction as $\vec{E}$. Equation (25.3) gives the *magnitude* J of the vector current density $\vec{J}$.

CAUTION Current density vs. current Current density $\vec{J}$ is a vector, but current I is not. The difference is that the current density $\vec{J}$ describes how charges flow at a certain point, and the vector's direction tells you about the direction of the flow at that point. By contrast, the current I describes how charges flow through an extended object such as a wire. For example, I has the same value at all points in the circuit of Fig. 25.3, but $\vec{J}$ does not: $\vec{J}$ is directed downward in the left-hand side of the loop and upward in the right-hand side. The magnitude of $\vec{J}$ can also vary around a circuit. In Fig. 25.3 $J = I/A$ is less in the battery (which has a large cross-sectional area A) than in the wires (which have a small A).

In general, a conductor may contain several different kinds of moving charged particles having charges $q_1, q_2, \ldots$, concentrations $n_1, n_2, \ldots$, and drift velocities with magnitudes $v_{\mathrm{d}1}, v_{\mathrm{d}2}, \ldots$. An example is current flow in an ionic solution (**Fig. 25.4**). In a sodium chloride solution, current can be carried by both positive sodium ions and negative chlorine ions; the total current I is found by adding up the currents due to each kind of charged particle, from Eq. (25.2). Likewise, the total vector current density $\vec{J}$ is found by using Eq. (25.4) for each kind of charged particle and adding the results.

We'll see in Section 25.4 that it is possible to have a current that is *steady* (constant in time) only if the conducting material forms a closed loop, called a *complete circuit*. In such a steady situation, the total charge in every segment of the conductor is constant. Hence the rate of flow of charge *out* at one end of a segment at any instant equals the rate of flow of charge *in* at the other end of the segment, and *the current is the same at all cross sections of the circuit*. We'll use this observation when we analyze electric circuits later in this chapter.

Figure **25.4** Part of the electric circuit that includes this light bulb passes through a beaker with a solution of sodium chloride. The current in the solution is carried by both positive charges (Na^+ ions) and negative charges (Cl^- ions).



In many simple circuits, such as flashlights or cordless electric drills, the direction of the current is always the same; this is called *direct current*. But home appliances such as toasters, refrigerators, and televisions use *alternating current,* in which the current continuously changes direction. In this chapter we'll consider direct current only. Alternating current has many special features worthy of detailed study, which we'll examine in Chapter 31.

EXAMPLE 25.1 Current density and drift velocity in a wire

An 18 gauge copper wire (the size usually used for lamp cords), with a diameter of 1.02 mm, carries a constant current of 1.67 A to a 200 W lamp. The free-electron density in the wire is 8.5×10^{28} per cubic meter. Find (a) the current density and (b) the drift speed.

IDENTIFY and SET UP This problem uses the relationships among current I, current density J, and drift speed v_d. We are given I and the wire diameter d, so we use Eq. (25.3) to find J. We use Eq. (25.3) again to find v_d from J and the known electron density n.

EXECUTE (a) The cross-sectional area is

$$A = \frac{\pi d^2}{4} = \frac{\pi(1.02 \times 10^{-3}\ \mathrm{m})^2}{4} = 8.17 \times 10^{-7}\ \mathrm{m}^2$$

The magnitude of the current density is then

$$J = \frac{I}{A} = \frac{1.67\ \mathrm{A}}{8.17 \times 10^{-7}\ \mathrm{m}^2} = 2.04 \times 10^6\ \mathrm{A/m^2}$$

(b) From Eq. (25.3) for the drift velocity magnitude v_d, we find

$$v_d = \frac{J}{n|q|} = \frac{2.04 \times 10^6\ \mathrm{A/m^2}}{(8.5 \times 10^{28}\ \mathrm{m^{-3}})|-1.60 \times 10^{-19}\ \mathrm{C}|}$$
$$= 1.5 \times 10^{-4}\ \mathrm{m/s} = 0.15\ \mathrm{mm/s}$$

EVALUATE At this speed an electron would require 6700 s (almost 2 h) to travel 1 m along this wire. The speeds of random motion of the electrons are roughly 10^6 m/s, around 10^{10} times the drift speed. Picture the electrons as bouncing around frantically, with a very slow drift!

KEYCONCEPT Current is the rate at which electric charge flows through an area, and current density is the current per unit area. Current density is proportional to the concentration of moving charged particles, the charge per particle, and the drift speed of the particles.

TEST YOUR UNDERSTANDING OF SECTION 25.1 Suppose we replaced the wire in Example 25.1 with 12 gauge copper wire, which has twice the diameter of 18 gauge wire. If the current remains the same, what effect would this have on the drift speed v_d? (i) None—v_d would be unchanged; (ii) v_d would be twice as great; (iii) v_d would be four times greater; (iv) v_d would be half as great; (v) v_d would be one-fourth as great.

ANSWER

(v) Doubling the diameter increases the cross-sectional area A by a factor of 4. Hence the current-density magnitude $J = I/A$ is reduced to $\frac{1}{4}$ of the value in Example 25.1, and the magnitude of the drift velocity $v_d = J/n|q|$ is reduced by the same factor. The new magnitude is $v_d = (0.15\ \mathrm{mm/s})/4 = 0.038\ \mathrm{mm/s}$. This behavior is the same as that of an incompressible fluid, which slows down when it moves from a narrow pipe to a broader one (see Section 14.4).

25.2 RESISTIVITY

The current density $\vec{J}$ in a conductor depends on the electric field $\vec{E}$ and on the properties of the material. In general, this dependence can be quite complex. But for some materials, especially metals, at a given temperature, $\vec{J}$ is nearly *directly proportional* to $\vec{E}$, and the ratio of the magnitudes of E and J is constant. This relationship, called Ohm's law, was discovered in 1826 by the German physicist Georg Simon Ohm (1787–1854). The word "law" should actually be in quotation marks, since **Ohm's law,** like the ideal-gas equation and Hooke's law, is an *idealized model* that describes the behavior of some materials quite well but is not a general description of *all* matter. In the following discussion we'll assume that Ohm's law is valid, even though there are many situations in which it is not.

We define the **resistivity** ρ of a material as

Resistivity of a material

$$\rho = \frac{E}{J} \qquad (25.5)$$

Magnitude of electric field in material

Magnitude of current density caused by electric field

TABLE 25.1 Resistivities at Room Temperature (20°C)

	Substance	$\rho\ (\Omega \cdot \mathrm{m})$	Substance	$\rho\ (\Omega \cdot \mathrm{m})$
Conductors			**Semiconductors**	
Metals	Silver	1.47×10^{-8}	Pure carbon (graphite)	3.5×10^{-5}
	Copper	1.72×10^{-8}	Pure germanium	0.60
	Gold	2.44×10^{-8}	Pure silicon	2300
	Aluminum	2.75×10^{-8}	**Insulators**	
	Tungsten	5.25×10^{-8}	Amber	5×10^{14}
	Steel	20×10^{-8}	Glass	10^{10}–10^{14}
	Lead	22×10^{-8}	Lucite	$>10^{13}$
	Mercury	95×10^{-8}	Mica	10^{11}–10^{15}
Alloys	Manganin (Cu 84%, Mn 12%, Ni 4%)	44×10^{-8}	Quartz (fused)	75×10^{16}
	Constantan (Cu 60%, Ni 40%)	49×10^{-8}	Sulfur	10^{15}
	Nichrome	100×10^{-8}	Teflon	$>10^{13}$
			Wood	10^{8}–10^{11}

The greater the resistivity, the greater the field needed to cause a given current density, or the smaller the current density caused by a given field. From Eq. (25.5) the units of ρ are $(\mathrm{V/m})/(\mathrm{A/m^2}) = \mathrm{V \cdot m/A}$. As we'll discuss in Section 25.3, 1 V/A is called one *ohm* (1 Ω; the Greek letter Ω, omega, is alliterative with "ohm"). So the SI units for ρ are $\Omega \cdot \mathrm{m}$ (ohm-meters). **Table 25.1** lists some representative values of resistivity. A perfect conductor would have zero resistivity, and a perfect insulator would have an infinite resistivity. Metals and alloys have the smallest resistivities and are the best conductors. The resistivities of insulators are greater than those of the metals by an enormous factor, on the order of 10^{22}.

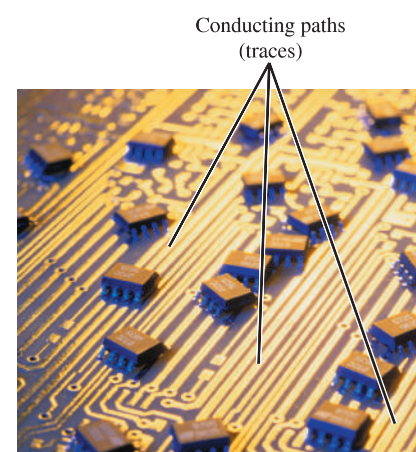
The reciprocal of resistivity is **conductivity.** Its units are $(\Omega \cdot \mathrm{m})^{-1}$. Good conductors of electricity have larger conductivity than insulators. Conductivity is the direct electrical analog of thermal conductivity. Comparing Table 25.1 with Table 17.5 (Thermal Conductivities), we note that good electrical conductors, such as metals, are usually also good conductors of heat. Poor electrical conductors, such as ceramic and plastic materials, are also poor thermal conductors. In a metal the free electrons that carry charge in electrical conduction also provide the principal mechanism for heat conduction, so we should expect a correlation between electrical and thermal conductivity. Because of the enormous difference in conductivity between electrical conductors and insulators, it is easy to confine electric currents to well-defined paths or circuits (**Fig. 25.5**). The variation in *thermal* conductivity is much less, only a factor of 10^3 or so, and it is usually impossible to confine heat currents to that extent.

Semiconductors have resistivities intermediate between those of metals and those of insulators. These materials are important because of the way their resistivities are affected by temperature and by small amounts of impurities.

A material that obeys Ohm's law reasonably well is called an *ohmic* conductor or a *linear* conductor. For such materials, at a given temperature, ρ is a *constant* that does not depend on the value of E. Many materials show substantial departures from Ohm's-law behavior; they are *nonohmic,* or *nonlinear.* In these materials, J depends on E in a more complicated manner.

Analogies with fluid flow can be a big help in developing intuition about electric current and circuits. For example, in the making of wine or maple syrup, the product is sometimes filtered to remove sediments. A pump forces the fluid through the filter under pressure; if the flow rate (analogous to J) is proportional to the pressure difference between the upstream and downstream sides (analogous to E), the behavior is analogous to Ohm's law.

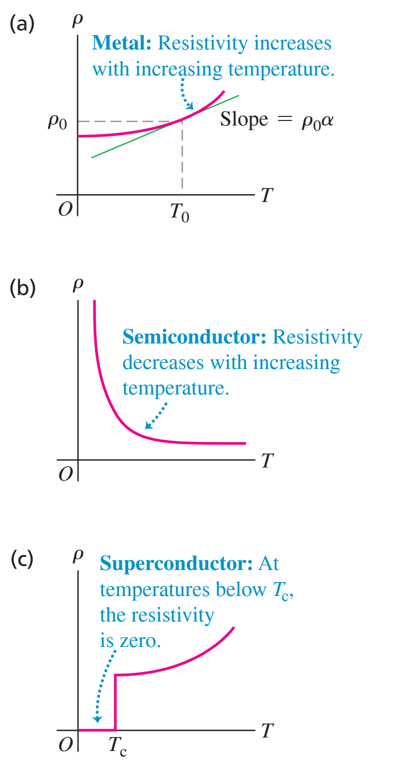
Figure **25.5** The copper "wires," or traces, on this circuit board are printed directly onto the surface of the dark-colored insulating board. Even though the traces are very close to each other (only about a millimeter apart), the board has such a high resistivity (and low conductivity) that essentially no current can flow between the traces.



Resistivity and Temperature

The resistivity of a *metallic* conductor nearly always increases with increasing temperature, as shown in **Fig. 25.6a** (next page). As temperature increases, the ions of the conductor vibrate with greater amplitude, making it more likely that a moving electron will

Figure **25.6** Variation of resistivity ρ with absolute temperature T for (a) a normal metal, (b) a semiconductor, and (c) a superconductor. In (a) the linear approximation to ρ as a function of T is shown as a green line; the approximation agrees exactly at $T = T_0$, where $\rho = \rho_0$.



collide with an ion as in Fig. 25.1; this impedes the drift of electrons through the conductor and hence reduces the current. Over a small temperature range (up to 100 C° or so), the resistivity of a metal can be represented approximately by the equation

Temperature dependence of resistivity:

Resistivity at temperature T · Temperature coefficient of resistivity

$$\rho(T) = \rho_0[1 + \alpha(T - T_0)] \qquad (25.6)$$

Resistivity at reference temperature T_0

The reference temperature T_0 is often taken as 0°C or 20°C; the temperature T may be higher or lower than T_0. The factor α is called the **temperature coefficient of resistivity.** Some representative values are given in **Table 25.2**. The resistivity of the alloy manganin is practically independent of temperature.

The resistivity of graphite (a nonmetal) *decreases* with increasing temperature, since at higher temperatures, more electrons "shake loose" from the atoms and become mobile; hence the temperature coefficient of resistivity of graphite is negative. The same behavior occurs for semiconductors (Fig. 25.6b). Measuring the resistivity of a small semiconductor crystal is therefore a sensitive measure of temperature; this is the principle of a type of thermometer called a *thermistor*.

Some materials, including several metallic alloys and oxides, show a phenomenon called *superconductivity*. As the temperature decreases, the resistivity at first decreases smoothly, like that of any metal. But then at a certain critical temperature T_c a phase transition occurs and the resistivity suddenly drops to zero, as shown in Fig. 25.6c. Once a current has been established in a superconducting ring, it continues indefinitely without the presence of any driving field.

Superconductivity was discovered in 1911 by the Dutch physicist Heike Kamerlingh Onnes (1853–1926). He discovered that at very low temperatures, below 4.2 K, the resistivity of mercury suddenly dropped to zero. For the next 75 years, the highest T_c attained was about 20 K. This meant that superconductivity occurred only when the material was cooled by using expensive liquid helium, with a boiling-point temperature of 4.2 K, or explosive liquid hydrogen, with a boiling point of 20.3 K. But in 1986 Karl Müller and Johannes Bednorz discovered an oxide of barium, lanthanum, and copper with a T_c of nearly 40 K, and the race was on to develop "high-temperature" superconducting materials.

By 1987 a complex oxide of yttrium, copper, and barium had been found that has a value of T_c well above the 77 K boiling temperature of liquid nitrogen, a refrigerant that is both inexpensive and safe. The current (2018) record for T_c is 203.5 K, and it may be possible to fabricate materials that are superconductors at room temperature. The implications of these discoveries for power-distribution systems, computer design, and transportation are enormous. Meanwhile, superconducting electromagnets cooled by liquid helium are used in particle accelerators and some experimental magnetic-levitation railroads. Superconductors have other exotic properties that require an understanding of magnetism to explore; we'll discuss these further in Chapter 29.

BIO APPLICATION Resistivity and Nerve Conduction This false-color image from an electron microscope shows a cross section through a nerve fiber about 1 μm (10^{-6} m) in diameter. A layer of an insulating fatty substance called myelin is wrapped around the conductive material of the axon. The resistivity of myelin is much greater than that of the axon, so an electric signal traveling along the nerve fiber remains confined to the axon. This makes it possible for a signal to travel much more rapidly than if the myelin were absent.

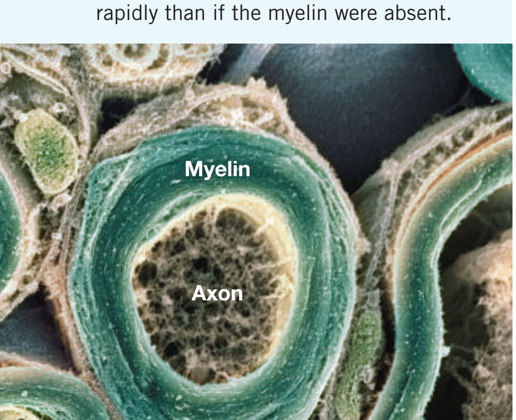


TABLE 25.2 Temperature Coefficients of Resistivity (Approximate Values Near Room Temperature)

Material	α [(°C)$^{-1}$]	Material	α [(°C)$^{-1}$]
Aluminum	0.0039	Lead	0.0043
Brass	0.0020	Manganin	0.00000
Carbon (graphite)	−0.0005	Mercury	0.00088
Constantan	0.000008	Nichrome	0.0004
Copper	0.00393	Silver	0.0038
Iron	0.0050	Tungsten	0.0045

TEST YOUR UNDERSTANDING OF SECTION 25.2 You maintain a constant electric field inside a piece of semiconductor while lowering the semiconductor's temperature. What happens to the current density in the semiconductor? (i) It increases; (ii) it decreases; (iii) it remains the same.

ANSWER

(ii) Figure 25.6b shows that the resistivity ρ of a semiconductor increases as the temperature decreases. From Eq. (25.5), the magnitude of the current density is $J = E/\rho$, so the current density decreases as the temperature drops and the resistivity increases.

25.3 RESISTANCE

For a conductor with resistivity ρ, the current density $\vec{J}$ at a point where the electric field is $\vec{E}$ is given by Eq. (25.5), which we can write as

$$\vec{E} = \rho\vec{J} \tag{25.7}$$

When Ohm's law is obeyed, ρ is constant and independent of the magnitude of the electric field, so $\vec{E}$ is directly proportional to $\vec{J}$. Often, however, we are more interested in the total current I in a conductor than in $\vec{J}$ and more interested in the potential difference V between the ends of the conductor than in $\vec{E}$. This is so largely because I and V are much easier to measure than are $\vec{J}$ and $\vec{E}$.

Suppose our conductor is a wire with uniform cross-sectional area A and length L, as shown in **Fig. 25.7**. Let V be the potential difference between the higher-potential and lower-potential ends of the conductor, so that V is positive. (Another name for V is the voltage across the conductor.) The *direction* of the current is always from the higher-potential end to the lower-potential end. That's because current in a conductor flows in the direction of $\vec{E}$, no matter what the sign of the moving charges (Fig. 25.2), and because $\vec{E}$ points in the direction of *decreasing* electric potential (see Section 23.2). As the current flows through the potential difference, electric potential energy is lost; this energy is transferred to the ions of the conducting material during collisions.

We can also relate the *value* of the current I to the potential difference between the ends of the conductor. If the magnitudes of the current density $\vec{J}$ and the electric field $\vec{E}$ are uniform throughout the conductor, the total current I is $I = JA$, and the potential difference V between the ends is $V = EL$. We solve these equations for J and E, respectively, and substitute the results into Eq. (25.7):

$$\frac{V}{L} = \frac{\rho I}{A} \quad \text{or} \quad V = \frac{\rho L}{A}I \tag{25.8}$$

This shows that when ρ is constant, the total current I is proportional to the potential difference V.

The ratio of V to I for a particular conductor is called its **resistance** R:

$$R = \frac{V}{I} \tag{25.9}$$

Comparing this definition of R to Eq. (25.8), we see that

Resistance of a conductor

$$R = \frac{\rho L}{A} \tag{25.10}$$

Resistivity of conductor material
Length of conductor
Cross-sectional area of conductor

If ρ is constant, as is the case for ohmic materials, then so is R.

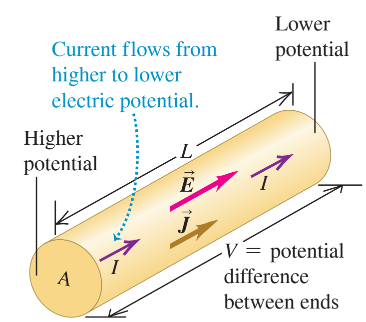
The following equation is often called Ohm's law:

Relationship among voltage, current, and resistance:

$$V = IR \tag{25.11}$$

Voltage between ends of conductor
Resistance of conductor
Current in conductor

Figure **25.7** A conductor with uniform cross section. The current density is uniform over any cross section, and the electric field is constant along the length.



However, it's important to understand that the real content of Ohm's law is the direct proportionality (for some materials) of V to I or of J to E. Equation (25.9) or (25.11) *defines* resistance R for *any* conductor, but only when R is constant can we correctly call this relationship Ohm's law.

Interpreting Resistance

Equation (25.10) shows that the resistance of a wire or other conductor of uniform cross section is directly proportional to its length and inversely proportional to its cross-sectional area. It is also proportional to the resistivity of the material of which the conductor is made.

The flowing-fluid analogy is again useful. In analogy to Eq. (25.10), a narrow water hose offers more resistance to flow than a fat one, and a long hose has more resistance than a short one (**Fig. 25.8**). We can increase the resistance to flow by stuffing the hose with cotton or sand; this corresponds to increasing the resistivity. The flow rate is approximately proportional to the pressure difference between the ends. Flow rate is analogous to current, and pressure difference is analogous to potential difference (voltage). Let's not stretch this analogy too far, though; the water flow rate in a pipe is usually *not* proportional to its cross-sectional area (see Section 14.6).

Figure **25.8** A long fire hose offers substantial resistance to water flow. To make water pass through the hose rapidly, the upstream end of the hose must be at much higher pressure than the end where the water emerges. In an analogous way, there must be a large potential difference between the ends of a long wire in order to cause a substantial electric current through the wire.

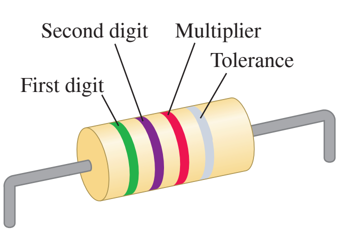
The SI unit of resistance is the **ohm,** equal to one volt per ampere $(1\ \Omega = 1\ \text{V/A})$. The *kilohm* $(1\ \text{k}\Omega = 10^3\ \Omega)$ and the *megohm* $(1\ \text{M}\Omega = 10^6\ \Omega)$ are also in common use. A 100 m length of 12 gauge copper wire, the size usually used in household wiring, has a resistance at room temperature of about $0.5\ \Omega$. A 100 W, 120 V incandescent light bulb has a resistance (at operating temperature) of $140\ \Omega$. If the same current I flows in both the copper wire and the light bulb, the potential difference $V = IR$ is much greater across the light bulb, and much more potential energy is lost per charge in the light bulb. This lost energy is converted by the light bulb filament into light and heat. You don't want your household wiring to glow white-hot, so its resistance is kept low by using wire of low resistivity and large cross-sectional area.

Because the resistivity of a material varies with temperature, the resistance of a specific conductor also varies with temperature. For temperature ranges that are not too great, this variation is approximately linear, analogous to Eq. (25.6):

$$R(T) = R_0[1 + \alpha(T - T_0)] \tag{25.12}$$

In this equation, $R(T)$ is the resistance at temperature T and R_0 is the resistance at temperature T_0, often taken to be 0°C or 20°C. The *temperature coefficient of resistance* α is the same constant that appears in Eq. (25.6) if the dimensions L and A in Eq. (25.10) do not change appreciably with temperature; this is indeed the case for most conducting materials. Within the limits of validity of Eq. (25.12), the *change* in resistance resulting from a temperature change $T - T_0$ is given by $R_0\alpha(T - T_0)$.

Figure **25.9** This resistor has a resistance of $5.7\ \text{k}\Omega$ with an accuracy (tolerance) of $\pm 10\%$.



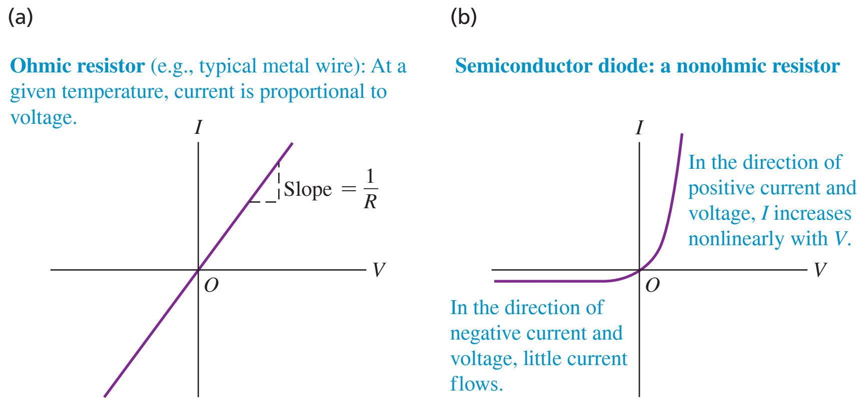
A circuit device made to have a specific value of resistance between its ends is called a **resistor.** Resistors in the range 0.01 to $10^7\ \Omega$ can be bought off the shelf. Individual resistors used in electronic circuitry are often cylindrical, a few millimeters in diameter and length, with wires coming out of the ends. The resistance may be marked with a standard code that uses three or four color bands near one end (**Fig. 25.9**), according to the scheme in **Table 25.3**. The first two bands (starting with the band nearest an end) are digits, and the third is a power-of-10 multiplier. For example, green–violet–red means $57 \times 10^2\ \Omega$, or $5.7\ \text{k}\Omega$. The fourth band, if present, indicates the accuracy (tolerance) of the value; no band means $\pm 20\%$, a silver band $\pm 10\%$, and a gold band $\pm 5\%$. Another important characteristic of a resistor is the maximum *power* it can dissipate without damage. We'll return to this point in Section 25.5.

TABLE 25.3 Color Codes for Resistors

Color	Value as Digit	Value as Multiplier
Black	0	1
Brown	1	10
Red	2	10^2
Orange	3	10^3
Yellow	4	10^4
Green	5	10^5
Blue	6	10^6
Violet	7	10^7
Gray	8	10^8
White	9	10^9

For a resistor that obeys Ohm's law, a graph of current as a function of potential difference (voltage) is a straight line (**Fig. 25.10a**). The slope of the line is $1/R$. If the sign of the potential difference changes, so does the sign of the current produced; in Fig. 25.7 this corresponds to interchanging the higher- and lower-potential ends of the conductor, so the electric field, current density, and current all reverse direction.

Figure 25.10 Current–voltage relationships for two devices. Only for a resistor that obeys Ohm's law as in (a) is current I proportional to voltage V.



In devices that do not obey Ohm's law, the relationship of voltage to current may not be a direct proportion, and it may be different for the two directions of current. Figure 25.10b shows the behavior of a semiconductor *diode,* a device used to convert alternating current to direct current and to perform a wide variety of logic functions in computer circuitry. For positive potentials V of the anode (one of two terminals of the diode) with respect to the cathode (the other terminal), I increases exponentially with increasing V; for negative potentials the current is extremely small. Thus a positive V causes a current to flow in the positive direction, but a potential difference of the other sign causes little or no current. Hence a diode acts like a one-way valve in a circuit.

EXAMPLE 25.2 Electric field, potential difference, and resistance in a wire

WITH VARIATION PROBLEMS

The 18 gauge copper wire of Example 25.1 has a cross-sectional area of $8.17 \times 10^{-7}\ \text{m}^2$. It carries a current of 1.67 A. Find (a) the electric-field magnitude in the wire; (b) the potential difference between two points in the wire 50.0 m apart; (c) the resistance of a 50.0 m length of this wire.

IDENTIFY and SET UP We are given the cross-sectional area A and current I. Our target variables are the electric-field magnitude E, potential difference V, and resistance R. The current density is $J = I/A$. We find E from Eq. (25.5), $E = \rho J$ (Table 25.1 gives the resistivity ρ for copper). The potential difference is then the product of E and the length of the wire. We can use either Eq. (25.10) or Eq. (25.11) to find R.

EXECUTE (a) From Table 25.1, $\rho = 1.72 \times 10^{-8}\ \Omega \cdot \text{m}$. Hence, from Eq. (25.5),

$$E = \rho J = \frac{\rho I}{A} = \frac{(1.72 \times 10^{-8}\ \Omega \cdot \text{m})(1.67\ \text{A})}{8.17 \times 10^{-7}\ \text{m}^2} = 0.0352\ \text{V/m}$$

(b) The potential difference is

$$V = EL = (0.0352\ \text{V/m})(50.0\ \text{m}) = 1.76\ \text{V}$$

(c) From Eq. (25.10) the resistance of 50.0 m of this wire is

$$R = \frac{\rho L}{A} = \frac{(1.72 \times 10^{-8}\ \Omega \cdot \text{m})(50.0\ \text{m})}{8.17 \times 10^{-7}\ \text{m}^2} = 1.05\ \Omega$$

Alternatively, we can find R from Eq. (25.11):

$$R = \frac{V}{I} = \frac{1.76\ \text{V}}{1.67\ \text{A}} = 1.05\ \Omega$$

EVALUATE We emphasize that the resistance of the wire is *defined* to be the ratio of voltage to current. If the wire is made of nonohmic material, then R is different for different values of V but is always given by $R = V/I$. Resistance is also always given by $R = \rho L/A$; if the material is nonohmic, ρ is not constant but depends on E (or, equivalently, on $V = EL$).

KEYCONCEPT *Resistivity* is a property of a material: The greater the resistivity, the greater the electric field that must be present in the material to produce a given current density. *Resistance* is a property of a circuit element or device: The greater the resistance, the greater the potential difference must be across the circuit element to produce a given current. Resistance depends on the dimensions of the circuit element as well as on the resistivity of the material of which it is made.

EXAMPLE 25.3 Temperature dependence of resistance

WITH VARIATION PROBLEMS

Suppose the resistance of a copper wire is 1.05 Ω at 20°C. Find the resistance at 0°C and 100°C.

IDENTIFY and SET UP We are given the resistance $R_0 = 1.05\ \Omega$ at a reference temperature $T_0 = 20°\text{C}$. We use Eq. (25.12) to find the resistances at $T = 0°\text{C}$ and $T = 100°\text{C}$ (our target variables), taking the temperature coefficient of resistivity from Table 25.2.

EXECUTE From Table 25.2, $\alpha = 0.00393\ (\text{C}°)^{-1}$ for copper. Then from Eq. (25.12),

$$\begin{aligned} R &= R_0[1 + \alpha(T - T_0)] \\ &= (1.05\ \Omega)\{1 + [0.00393\ (\text{C}°)^{-1}][0°\text{C} - 20°\text{C}]\} \\ &= 0.97\ \Omega \text{ at } T = 0°\text{C} \end{aligned}$$

Continued

$$R = (1.05\ \Omega)\{1 + [0.00393\ (\mathrm{C}°)^{-1}][100°\mathrm{C} - 20°\mathrm{C}]\}$$
$$= 1.38\ \Omega \text{ at } T = 100°\mathrm{C}$$

EVALUATE The resistance at 100°C is greater than that at 0°C by a factor of $(1.38\ \Omega)/(0.97\ \Omega) = 1.42$: Raising the temperature of copper wire from 0°C to 100°C increases its resistance by 42%. From Eq. (25.11), $V = IR$, this means that 42% more voltage is required to produce the same current at 100°C than at 0°C. Designers of electric circuits that must operate over a wide temperature range must take this substantial effect into account.

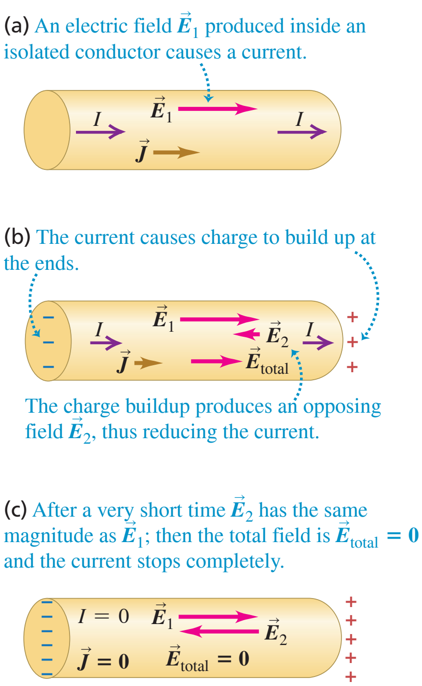
KEYCONCEPT The resistivity of a material depends on temperature, and so does the resistance of a circuit element made of that material. Over a small temperature range, the change in the resistivity of a metal is approximately proportional to the temperature change.

TEST YOUR UNDERSTANDING OF SECTION 25.3 Suppose you increase the voltage across the copper wire in Examples 25.2 and 25.3. The increased voltage causes more current to flow, which makes the temperature of the wire increase. (The same thing happens to the coils of an electric oven or a toaster when a voltage is applied to them. We'll explore this issue in more depth in Section 25.5.) If you double the voltage across the wire, the current in the wire increases. By what factor does it increase? (i) 2; (ii) greater than 2; (iii) less than 2.

ANSWER

(iii) Solving Eq. (25.11) for the current shows that $I = V/R$. If the resistance R of the wire remained the same, doubling the voltage V would make the current I double as well. However, we saw in Example 25.3 that the resistance is *not* constant: As the current increases and the temperature increases, R increases as well. Thus doubling the voltage produces a current that is *less* than double the original current. An ohmic conductor is one for which $R = V/I$ has the same value no matter what the voltage, so the wire is *nonohmic*. (In many practical problems the temperature change of the wire is so small that it can be ignored, so we can safely regard the wire as being ohmic. We do so in almost all examples in this book.)

Figure 25.11 If an electric field is produced inside a conductor that is *not* part of a complete circuit, current flows for only a very short time.



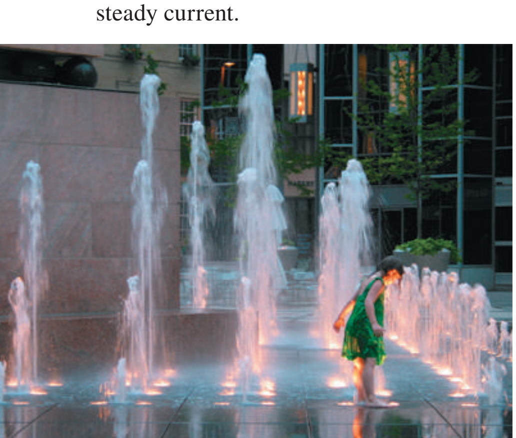
25.4 ELECTROMOTIVE FORCE AND CIRCUITS

For a conductor to have a steady current, it must be part of a path that forms a closed loop or **complete circuit.** Here's why. If you establish an electric field $\vec{E}_1$ inside an isolated conductor with resistivity ρ that is *not* part of a complete circuit, a current begins to flow with current density $\vec{J} = \vec{E}_1/\rho$ (**Fig. 25.11a**). As a result a net positive charge quickly accumulates at one end of the conductor and a net negative charge accumulates at the other end (Fig. 25.11b). These charges themselves produce an electric field $\vec{E}_2$ in the direction opposite to $\vec{E}_1$, causing the total electric field and hence the current to decrease. Within a very small fraction of a second, enough charge builds up on the conductor ends that the total electric field $\vec{E} = \vec{E}_1 + \vec{E}_2 = \mathbf{0}$ inside the conductor. Then $\vec{J} = \mathbf{0}$ as well, and the current stops altogether (Fig. 25.11c). So there can be no steady motion of charge in such an *incomplete* circuit.

To see how to maintain a steady current in a *complete* circuit, we recall a basic fact about electric potential energy: If a charge q goes around a complete circuit and returns to its starting point, the potential energy must be the same at the end of the round trip as at the beginning. As described in Section 25.3, there is always a *decrease* in potential energy when charges move through an ordinary conducting material with resistance. So there must be some part of the circuit in which the potential energy *increases.*

The problem is analogous to an ornamental water fountain that recycles its water. The water pours out of openings at the top, cascades down over the terraces and spouts (moving in the direction of decreasing gravitational potential energy), and collects in a basin in the bottom. A pump then lifts it back to the top (increasing the potential energy) for another trip. Without the pump, the water would just fall to the bottom and stay there.

Figure 25.12 Just as a water fountain requires a pump, an electric circuit requires a source of electromotive force to sustain a steady current.



Electromotive Force

In an electric circuit there must be a device somewhere in the loop that acts like the water pump in a water fountain (**Fig. 25.12**). In this device a charge travels "uphill," from lower to higher potential energy, even though the electrostatic force is trying to push it from higher to lower potential energy. The direction of current in such a device is from lower to higher potential, just the opposite of what happens in an ordinary conductor.

The influence that makes current flow from lower to higher potential is called **electromotive force** (abbreviated **emf** and pronounced "ee-em-eff"), and a circuit device that provides emf is called a **source of emf.** Note that "electromotive force" is a poor term because emf is *not* a force but an energy-per-unit-charge quantity, like potential. The SI unit of emf is the same as that for potential, the volt $(1\ \text{V} = 1\ \text{J/C})$. A typical flashlight battery has an emf of 1.5 V; this means that the battery does 1.5 J of work on every coulomb of charge that passes through it. We'll use the symbol $\mathcal{E}$ (a script capital E) for emf.

Every complete circuit with a steady current must include a source of emf. Batteries, electric generators, solar cells, thermocouples, and fuel cells are all examples of sources of emf. All such devices convert energy of some form (mechanical, chemical, thermal, and so on) into electric potential energy and transfer it into the circuit to which the device is connected. An *ideal* source of emf maintains a constant potential difference between its terminals, independent of the current through it. We define electromotive force quantitatively as the magnitude of this potential difference. As we'll see, such an ideal source is a mythical beast, like the frictionless plane and the massless rope. We'll discuss later how real-life sources of emf differ in their behavior from this idealized model.

Figure 25.13 is a schematic diagram of an ideal source of emf that maintains a potential difference between conductors a and b, called the *terminals* of the device. Terminal a, marked $+$, is maintained at *higher* potential than terminal b, marked $-$. Associated with this potential difference is an electric field $\vec{E}$ in the region around the terminals, both inside and outside the source. The electric field inside the device is directed from a to b, as shown. A charge q within the source experiences an electric force $\vec{F}_e = q\vec{E}$. But the source also provides an additional influence, which we represent as a nonelectrostatic force $\vec{F}_n$. This force, operating inside the device, pushes charge from b to a in an "uphill" direction against the electric force $\vec{F}_e$. Thus $\vec{F}_n$ maintains the potential difference between the terminals. If $\vec{F}_n$ were not present, charge would flow between the terminals until the potential difference was zero. The origin of the additional influence $\vec{F}_n$ depends on the kind of source. In a generator it results from magnetic-field forces on moving charges. In a battery or fuel cell it is associated with diffusion processes and varying electrolyte concentrations resulting from chemical reactions. In an electrostatic machine such as a Van de Graaff generator (see Fig. 22.26), an actual mechanical force is applied by a moving belt or wheel.

If a positive charge q is moved from b to a inside the source, the nonelectrostatic force $\vec{F}_n$ does a positive amount of work $W_n = q\mathcal{E}$ on the charge. This displacement is *opposite* to the electrostatic force $\vec{F}_e$, so the potential energy associated with the charge *increases* by an amount equal to qV_{ab}, where $V_{ab} = V_a - V_b$ is the (positive) potential of point a with respect to point b. For the ideal source of emf that we've described, $\vec{F}_e$ and $\vec{F}_n$ are equal in magnitude but opposite in direction, so the total work done on the charge q is zero; there is an increase in potential energy but *no* change in the kinetic energy of the charge. It's like lifting a book from the floor to a high shelf at constant speed. The increase in potential energy is just equal to the nonelectrostatic work W_n, so $q\mathcal{E} = qV_{ab}$, or

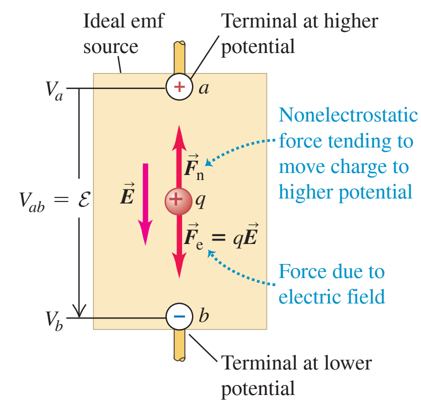
$$V_{ab} = \mathcal{E} \qquad \text{(ideal source of emf)} \qquad (25.13)$$

Now let's make a complete circuit by connecting a wire with resistance R to the terminals of a source (**Fig. 25.14**). The potential difference between terminals a and b sets up an electric field within the wire; this causes current to flow around the loop from a toward b, from higher to lower potential. Where the wire bends, equal amounts of positive and negative charges persist on the "inside" and "outside" of the bend. These charges exert the forces that cause the current to follow the bends in the wire.

From Eq. (25.11) the potential difference between the ends of the wire in Fig. 25.14 is given by $V_{ab} = IR$. Combining with Eq. (25.13), we have

$$\mathcal{E} = V_{ab} = IR \qquad \text{(ideal source of emf)} \qquad (25.14)$$

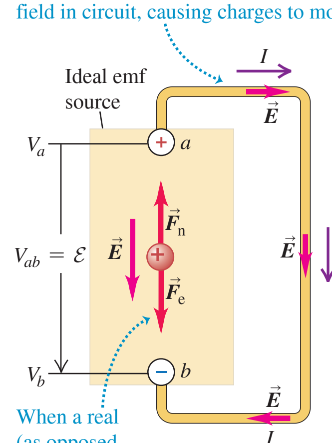
Figure **25.13** Schematic diagram of a source of emf in an "open-circuit" situation. The electric-field force $\vec{F}_e = q\vec{E}$ and the nonelectrostatic force $\vec{F}_n$ are shown for a positive charge q.



When the emf source is not part of a closed circuit, $F_n = F_e$ and there is no net motion of charge between the terminals.

Figure **25.14** Schematic diagram of an ideal source of emf in a complete circuit. The electric-field force $\vec{F}_e = q\vec{E}$ and the nonelectrostatic force $\vec{F}_n$ are shown for a positive charge q. The current is in the direction from a to b in the external circuit and from b to a within the source.

Potential across terminals creates electric field in circuit, causing charges to move.



When a real (as opposed to ideal) emf source is connected to a circuit, V_{ab} and thus F_e fall, so $F_n > F_e$ and $\vec{F}_n$ does work on the charges.

BIO APPLICATION Danger: Electric Ray! Electric rays deliver electric shocks to stun their prey and to discourage predators. (In ancient Rome, physicians practiced a primitive form of electroconvulsive therapy by placing electric rays on their patients to cure headaches and gout.) The shocks are produced by specialized flattened cells called electroplaques. Such a cell moves ions across membranes to produce an emf of about 0.05 V. Thousands of electroplaques are stacked on top of each other, so their emfs add to a total of as much as 200 V. These stacks make up more than half of an electric ray's body mass. A ray can use these to deliver an impressive current of up to 30 A for a few milliseconds.

That is, when a positive charge q flows around the circuit, the potential *rise* $\mathcal{E}$ as it passes through the ideal source is numerically equal to the potential *drop* $V_{ab} = IR$ as it passes through the remainder of the circuit. Once $\mathcal{E}$ and R are known, this relationship determines the current in the circuit.

CAUTION Current is not "used up" in a circuit It's a common misconception that in a closed circuit, current squirts out of the positive terminal of a battery and is consumed or "used up" by the time it reaches the negative terminal. In fact the current is the *same* at every point in a simple loop circuit like that in Fig. 25.14, even if the wire thickness is not constant throughout the circuit. This happens because charge is conserved (it can be neither created nor destroyed) and because charge cannot accumulate in the circuit devices we have described. It's like the flow of water in an ornamental fountain; water flows out of the top at the same rate at which it reaches the bottom, no matter what the dimensions of the fountain. None of the water is "used up" along the way!

Internal Resistance

Real sources of emf in a circuit don't behave in exactly the way we have described; the potential difference across a real source in a circuit is *not* equal to the emf as in Eq. (25.14). The reason is that charge moving through the material of any real source encounters *resistance*. We call this the **internal resistance** of the source, denoted by r. If this resistance behaves according to Ohm's law, r is constant and independent of the current I. As the current moves through r, it experiences an associated drop in potential equal to Ir. Thus, when a current is flowing through a source from the negative terminal b to the positive terminal a, the potential difference V_{ab} between the terminals is

Terminal voltage, source with internal resistance
emf of source
Current through source

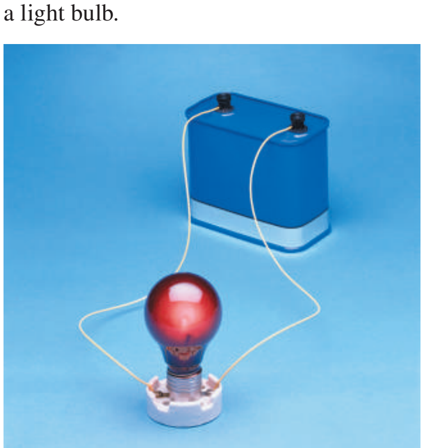
$$V_{ab} = \mathcal{E} - Ir \qquad (25.15)$$

Internal resistance of source

The potential V_{ab}, called the **terminal voltage,** is less than the emf $\mathcal{E}$ because of the term Ir representing the potential drop across the internal resistance r. Hence the increase in potential energy qV_{ab} as a charge q moves from b to a within the source is less than the work $q\mathcal{E}$ done by the nonelectrostatic force $\vec{F}_n$, since some potential energy is lost in traversing the internal resistance.

A 1.5 V battery has an emf of 1.5 V, but the terminal voltage V_{ab} of the battery is equal to 1.5 V only if no current is flowing through it so that $I = 0$ in Eq. (25.15). If the battery is part of a complete circuit through which current is flowing, the terminal voltage will be less than 1.5 V. *For a real source of emf, the terminal voltage equals the emf only if no current is flowing through the source* (**Fig. 25.15**). Thus we can describe the behavior of a source in terms of two properties: an emf $\mathcal{E}$, which supplies a constant potential difference independent of current, in series with an internal resistance r.

Figure 25.15 The emf of this battery—that is, the terminal voltage when it's not connected to anything—is 12 V. But because the battery has internal resistance, the terminal voltage of the battery is less than 12 V when it is supplying current to a light bulb.



The current in the external circuit connected to the source terminals a and b is still determined by $V_{ab} = IR$. Combining this with Eq. (25.15), we find

$$\mathcal{E} - Ir = IR \quad \text{or} \quad I = \frac{\mathcal{E}}{R + r} \quad \text{(current, source with internal resistance)} \qquad (25.16)$$

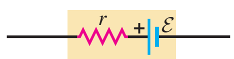
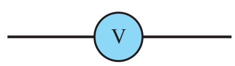
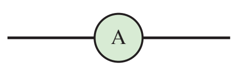
That is, the current equals the source emf divided by the *total* circuit resistance $(R + r)$.

CAUTION A battery is not a "current source" You might have thought that a battery or other source of emf always produces the same current, no matter what circuit it's used in. Equation (25.16) shows that this isn't so! The greater the resistance R of the external circuit, the less current the source will produce.

Symbols for Circuit Diagrams

An important part of analyzing any electric circuit is drawing a schematic *circuit diagram*. **Table 25.4** shows the usual symbols used in circuit diagrams. We'll use these symbols extensively in this chapter and the next. We usually assume that the wires

TABLE 25.4 Symbols for Circuit Diagrams

	Conductor with negligible resistance
	Resistor
	Source of emf (longer vertical line always represents the positive terminal, usually the terminal with higher potential)
 or 	Source of emf with internal resistance r (r can be placed on either side)
	Voltmeter (measures potential difference between its terminals)
	Ammeter (measures current through it)

CAUTION Internal resistance is an intrinsic part of a source of emf In Table 25.4 we've drawn the emf $\mathcal{E}$ and internal resistance r as distinct parts of a source of emf. In fact, both $\mathcal{E}$ and r are always present in any source. They *cannot* be separated.

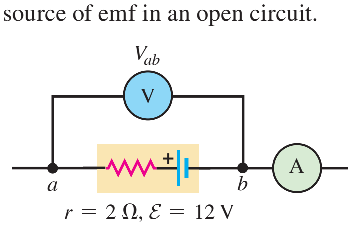
that connect the various elements of the circuit have negligible resistance; from Eq. (25.11), $V = IR$, the potential difference between the ends of such a wire is zero.

Table 25.4 includes two *meters* that are used to measure the properties of circuits. Idealized meters do not disturb the circuit in which they are connected. A **voltmeter,** introduced in Section 23.2, measures the potential difference between its terminals; an idealized voltmeter has infinitely large resistance and measures potential difference without having any current diverted through it. An ammeter measures the current passing through it; an idealized **ammeter** has zero resistance and has no potential difference between its terminals. The following examples illustrate how to analyze circuits that include meters.

CONCEPTUAL EXAMPLE 25.4 A source in an open circuit

Figure 25.16 shows a source (a battery) with emf $\mathcal{E} = 12$ V and internal resistance $r = 2\ \Omega$. (For comparison, the internal resistance of a commercial 12 V lead storage battery is only a few thousandths of an ohm.) The wires to the left of a and to the right of the ammeter A are not connected to anything. Determine the respective readings V_{ab} and I of the idealized voltmeter V and the idealized ammeter A.

Figure **25.16** A source of emf in an open circuit.



SOLUTION There is *zero* current because there is no complete circuit. (Our idealized voltmeter has an infinitely large resistance, so no current flows through it.) Hence the ammeter reads $I = 0$. Because there is no current through the battery, there is no potential difference across its internal resistance. From Eq. (25.15) with $I = 0$, the potential difference V_{ab} across the battery terminals is equal to the emf. So the voltmeter reads $V_{ab} = \mathcal{E} = 12$ V.

KEYCONCEPT The terminal voltage of a real, nonideal source with internal resistance equals the emf *only* if no current is flowing through the source.

EXAMPLE 25.5 A source in a complete circuit

WITH VARIATION PROBLEMS

We add a 4 Ω resistor to the battery in Conceptual Example 25.4, forming a complete circuit (**Fig. 25.17**, next page). What are the voltmeter and ammeter readings V_{ab} and I now?

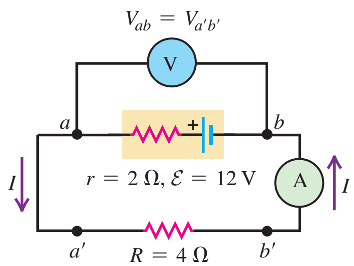
IDENTIFY and SET UP Our target variables are the current I through the circuit $aa'b'b$ and the potential difference V_{ab}. We first find I from Eq. (25.16). To find V_{ab}, we can use either Eq. (25.11) or Eq. (25.15).

EXECUTE The ideal ammeter has zero resistance, so the total resistance external to the source is $R = 4\ \Omega$. From Eq. (25.16), the current through the circuit $aa'b'b$ is then

$$I = \frac{\mathcal{E}}{R + r} = \frac{12\text{ V}}{4\ \Omega + 2\ \Omega} = 2\text{ A}$$

Continued

Figure **25.17** A source of emf in a complete circuit.



Our idealized conducting wires and the idealized ammeter have zero resistance, so there is no potential difference between points a and a' or between points b and b'; that is, $V_{ab} = V_{a'b'}$. We find V_{ab} by considering a and b as the terminals of the resistor: From Ohm's law, Eq. (25.11), we then have

$$V_{a'b'} = IR = (2\ \text{A})(4\ \Omega) = 8\ \text{V}$$

Alternatively, we can consider a and b as the terminals of the source. Then, from Eq. (25.15),

$$V_{ab} = \mathcal{E} - Ir = 12\ \text{V} - (2\ \text{A})(2\ \Omega) = 8\ \text{V}$$

Either way, we see that the voltmeter reading is 8 V.

EVALUATE In this example the battery is acting as a source that produces the current in the circuit. Current flows through the battery from its negative terminal toward its positive terminal, and the terminal voltage V_{ab} is less than the emf $\mathcal{E}$. But if a battery is being recharged, current flows through it in the opposite direction, from its *positive* terminal toward its *negative* terminal. In such a case I in Eq. (25.15) is negative, so the terminal voltage is actually *greater* than the emf. No matter which way current flows through the battery, the smaller the internal resistance r, the less the difference between V_{ab} and $\mathcal{E}$.

KEYCONCEPT To find the potential difference across a resistor, multiply the current through the resistor by its resistance. The potential difference across a source, such as a battery, differs from its emf by an amount determined by the source's internal resistance.

CONCEPTUAL EXAMPLE 25.6 Using voltmeters and ammeters

We move the voltmeter and ammeter in Example 25.5 to different positions in the circuit. What are the readings of the ideal voltmeter and ammeter in the situations shown in (a) **Fig. 25.18a** and (b) Fig. 25.18b?

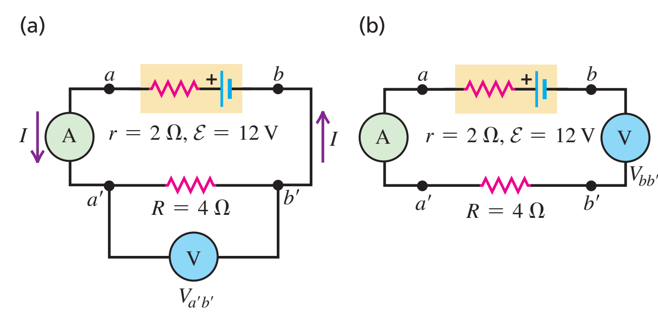
SOLUTION (a) The voltmeter now measures the potential difference between points a' and b'. As in Example 25.5, $V_{ab} = V_{a'b'}$, so the voltmeter reads the same as in Example 25.5: $V_{a'b'} = 8\ \text{V}$.

CAUTION Current in a simple loop As charges move through a resistor, there is a decrease in electric potential energy, but there is *no* change in the current. *The current in a simple loop is the same at every point;* it is not "used up" as it moves through a resistor. Hence the ammeter in Fig. 25.17 ("downstream" of the 4 Ω resistor) and the ammeter in Fig. 25.18a ("upstream" of the resistor) both read $I = 2$ A.

(b) There is no current through the ideal voltmeter because it has infinitely large resistance. Since the voltmeter is now part of the circuit, there is no current at all in the circuit, and the ammeter reads $I = 0$.

The voltmeter measures the potential difference $V_{bb'}$ between points b and b'. Since $I = 0$, the potential difference across the resistor is $V_{a'b'} = IR = 0$, and the potential difference between the ends a and a' of the idealized ammeter is also zero. So $V_{bb'}$ is equal to V_{ab}, the terminal voltage of the source. As in Conceptual Example 25.4, there is no current, so the terminal voltage equals the emf, and the voltmeter reading is $V_{ab} = \mathcal{E} = 12\ \text{V}$.

Figure **25.18** Different placements of a voltmeter and an ammeter in a complete circuit.



This example shows that ammeters and voltmeters are circuit elements, too. Moving the voltmeter from the position in Fig. 25.18a to that in Fig. 25.18b makes large changes in the current and potential differences in the circuit. If you want to measure the potential difference between two points in a circuit without disturbing the circuit, use a voltmeter as in Fig. 25.17 or 25.18a, *not* as in Fig. 25.18b.

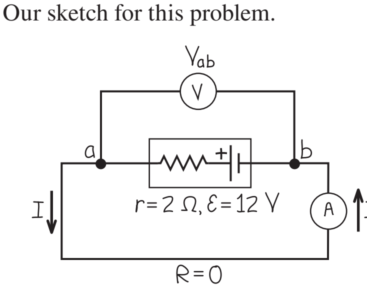
KEYCONCEPT An ideal voltmeter has infinite resistance, whereas an ideal ammeter has zero resistance. Placing these meters in inappropriate locations in a circuit can disturb the quantities you are trying to measure.

EXAMPLE 25.7 A source with a short circuit

In the circuit of Example 25.5 we replace the 4 Ω resistor with a zero-resistance conductor. What are the meter readings now?

IDENTIFY and SET UP **Figure 25.19** shows the new circuit. Our target variables are again I and V_{ab}. There is now a zero-resistance path between points a and b, through the lower loop, so the potential difference between these points must be zero.

Figure **25.19** Our sketch for this problem.



EXECUTE We must have $V_{ab} = IR = I(0) = 0$, no matter what the current. We can therefore find the current I from Eq. (25.15):

$$V_{ab} = \mathcal{E} - Ir = 0$$

$$I = \frac{\mathcal{E}}{r} = \frac{12\ \text{V}}{2\ \Omega} = 6\ \text{A}$$

EVALUATE The current in this circuit has a different value than in Example 25.5, even though the same battery is used; the current depends on both the internal resistance r and the resistance of the external circuit.

The situation here is called a *short circuit*. The external-circuit resistance is zero, because terminals of the battery are connected directly to each other. The short-circuit current is equal to the emf $\mathcal{E}$ divided by the internal resistance r. *Warning:* Short circuits can be dangerous! An automobile battery or a household power line has very small internal resistance (much less than in these examples), and the short-circuit current can be great enough to melt a small wire or cause a storage battery to explode.

KEYCONCEPT Connecting the terminals of a source directly to each other (a short circuit) leads to a large and potentially unsafe current.

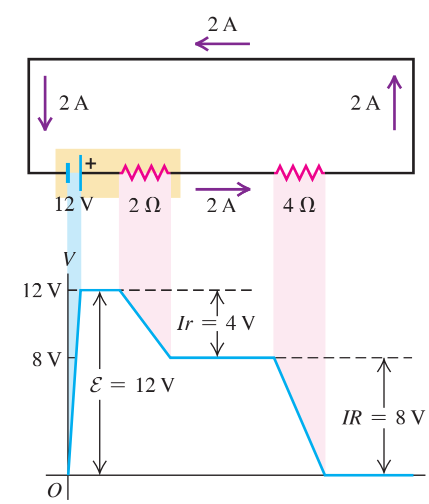
Potential Changes Around a Circuit

The net change in potential energy for a charge q making a round trip around a complete circuit must be zero. Hence the net change in *potential* around the circuit must also be zero; in other words, the algebraic sum of the potential differences and emfs around the loop is zero. We can see this by rewriting Eq. (25.16) in the form

$$\mathcal{E} - Ir - IR = 0$$

A potential gain of $\mathcal{E}$ is associated with the emf, and potential drops of Ir and IR are associated with the internal resistance of the source and the external circuit, respectively. **Figure 25.20** shows how the potential varies as we go around the complete circuit of Fig. 25.17. The horizontal axis doesn't necessarily represent actual distances, but rather various points in the loop. If we take the potential to be zero at the negative terminal of the battery, then we have a rise $\mathcal{E}$ and a drop Ir in the battery and an additional drop IR in the external resistor, and as we finish our trip around the loop, the potential is back where it started.

Figure 25.20 Potential rises and drops in a circuit.



In this section we have considered only situations in which the resistances are ohmic. If the circuit includes a nonlinear device such as a diode (see Fig. 25.10b), Eq. (25.16) is still valid but cannot be solved algebraically because R is not a constant. In such a situation, I can be found by using numerical techniques.

Finally, we remark that Eq. (25.15) is not always an adequate representation of the behavior of a source. The emf may not be constant, and what we have described as an internal resistance may actually be a more complex voltage–current relationship that doesn't obey Ohm's law. Nevertheless, the concept of internal resistance frequently provides an adequate description of batteries, generators, and other energy converters. The principal difference between a fresh flashlight battery and an old one is not in the emf, which decreases only slightly with use, but in the internal resistance, which may increase from less than an ohm when the battery is fresh to as much as 1000 Ω or more after long use. Similarly, a car battery can deliver less current to the starter motor on a cold morning than when the battery is warm, not because the emf is appreciably less but because the internal resistance increases with decreasing temperature.

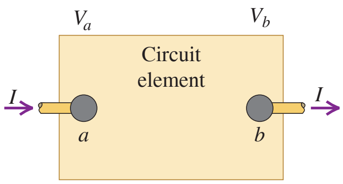
TEST YOUR UNDERSTANDING OF SECTION 25.4 Rank the following circuits in order from highest to lowest current: (i) A 1.4 Ω resistor connected to a 1.5 V battery that has an internal resistance of 0.10 Ω; (ii) a 1.8 Ω resistor connected to a 4.0 V battery that has a terminal voltage of 3.6 V but an unknown internal resistance; (iii) an unknown resistor connected to a 12.0 V battery that has an internal resistance of 0.20 Ω and a terminal voltage of 11.0 V.

ANSWER

(iii), (ii), (i) For circuit (i), we find the current from Eq. (25.16): $I = \mathcal{E}/(R + r) = (1.5\ \text{V})/(1.4\ \Omega + 0.10\ \Omega) = 1.0\ \text{A}$. For circuit (ii), we note that the terminal voltage $V_{ab} = 3.6\ \text{V}$ equals the voltage IR across the 1.8 Ω resistor: $V_{ab} = IR$, so $I = V_{ab}/R = (3.6\ \text{V})/(1.8\ \Omega) = 2.0\ \text{A}$. For circuit (iii), we use Eq. (25.15) for the terminal voltage: $V_{ab} = \mathcal{E} - Ir$, so $I = (\mathcal{E} - V_{ab})/r = (12.0\ \text{V} - 11.0\ \text{V})/(0.20\ \Omega) = 5.0\ \text{A}$.

25.5 ENERGY AND POWER IN ELECTRIC CIRCUITS

Figure **25.21** The power input to the circuit element between a and b is $P = (V_a - V_b)I = V_{ab}I$.



Let's now look at some energy and power relationships in electric circuits. The box in **Fig. 25.21** represents a circuit element with potential difference $V_a - V_b = V_{ab}$ between its terminals and current I passing through it in the direction from a toward b. This element might be a resistor, a battery, or something else; the details don't matter. As charge passes through the circuit element, the electric field does work on the charge. In a source of emf, additional work is done by the force $\vec{F}_{\text{n}}$ that we mentioned in Section 25.4.

As an amount of charge q passes through the circuit element, there is a change in potential energy equal to qV_{ab}. For example, if $q > 0$ and $V_{ab} = V_a - V_b$ is positive, potential energy decreases as the charge "falls" from potential V_a to lower potential V_b. The moving charges don't gain *kinetic* energy, because the current (the rate of charge flow) out of the circuit element must be the same as the current into the element. Instead, the quantity qV_{ab} represents energy transferred into the circuit element. This situation occurs in the coils of a toaster or electric oven, in which electrical energy is converted to thermal energy.

If the potential at a is lower than at b, then V_{ab} is negative and there is a net transfer of energy *out* of the circuit element. The element then acts as a source, delivering electrical energy into the circuit to which it is attached. This is the usual situation for a battery, which converts chemical energy into electrical energy and delivers it to the external circuit. Thus qV_{ab} can denote a quantity of energy that is either delivered to a circuit element or extracted from that element.

In electric circuits we are most often interested in the *rate* at which energy is either delivered to or extracted from a circuit element. If the current through the element is I, then in a time interval dt an amount of charge $dQ = I\,dt$ passes through the element. The potential-energy change for this amount of charge is $V_{ab}\,dQ = V_{ab}I\,dt$. Dividing this expression by dt, we obtain the *rate* at which energy is transferred either into or out of the circuit element. The time rate of energy transfer is *power,* denoted by P, so we write

$$\underset{\text{Power delivered to or extracted from a circuit element}}{P} = \overset{\text{Voltage across circuit element}}{V_{ab}}\,\underset{\text{Current in circuit element}}{I} \qquad (25.17)$$

The unit of V_{ab} is one volt, or one joule per coulomb, and the unit of I is one ampere, or one coulomb per second. Hence the unit of $P = V_{ab}I$ is one watt:

$$(1\ \text{J/C})(1\ \text{C/s}) = 1\ \text{J/s} = 1\ \text{W}$$

Let's consider a few special cases.

Power Input to a Pure Resistance

If the circuit element in Fig. 25.21 is a resistor, the potential difference is $V_{ab} = IR$. From Eq. (25.17) the power delivered to the resistor by the circuit is

$$\underset{\text{Power delivered to a resistor}}{P} = \overset{\text{Voltage across resistor}}{V_{ab}}\,\underset{\text{Current in resistor}}{I} = I^2R = \frac{V_{ab}^{\,2}}{\underset{\text{Resistance of resistor}}{R}} \qquad (25.18)$$

In this case the potential at a (where the current enters the resistor) is always higher than that at b (where the current exits). Current enters the higher-potential terminal of the device, and Eq. (25.18) represents the rate of transfer of electric potential energy *into* the circuit element.

What becomes of this energy? The moving charges collide with atoms in the resistor and transfer some of their energy to these atoms, increasing the internal energy of the material. Either the temperature of the resistor increases or there is a flow of heat out of it, or both. In any of these cases we say that energy is *dissipated* in the resistor at a rate I^2R. Every resistor has a *power rating,* the maximum power the device can dissipate without becoming overheated and damaged. Some devices, such as electric heaters, are designed to get hot and transfer heat to their surroundings. But if the power rating is exceeded, even such a device may melt or even explode.

Power Output of a Source

The upper rectangle in **Fig. 25.22a** represents a source with emf $\mathcal{E}$ and internal resistance r, connected by ideal (resistanceless) conductors to an external circuit represented by the lower box. This could describe a car battery connected to one of the car's headlights (Fig. 25.22b). Point a is at higher potential than point b, so $V_a > V_b$ and V_{ab} is positive. Note that the current I is *leaving* the source at the higher-potential terminal (rather than entering there). Energy is being delivered to the external circuit, at a rate given by Eq. (25.17):

$$P = V_{ab}I$$

For a source that can be described by an emf $\mathcal{E}$ and an internal resistance r, we may use Eq. (25.15):

$$V_{ab} = \mathcal{E} - Ir$$

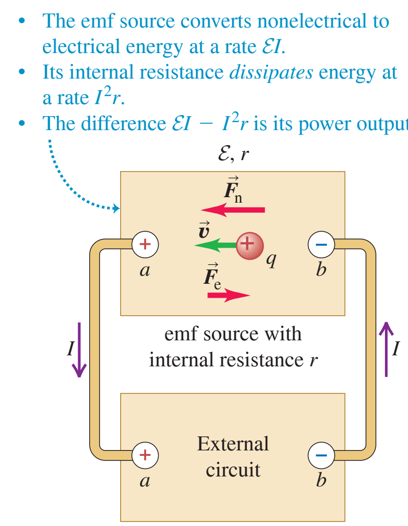
Multiplying this equation by I, we find

$$P = V_{ab}I = \mathcal{E}I - I^2r \tag{25.19}$$

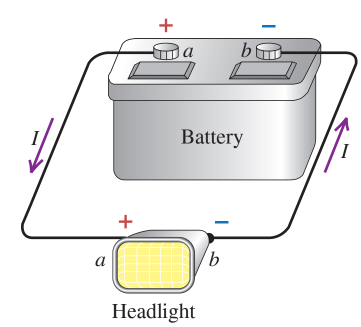
What do the terms $\mathcal{E}I$ and I^2r mean? In Section 25.4 we defined the emf $\mathcal{E}$ as the work per unit charge performed on the charges by the nonelectrostatic force as the charges are pushed "uphill" from b to a in the source. In a time dt, a charge $dQ = I\,dt$ flows through the source; the work done on it by this nonelectrostatic force is $\mathcal{E}\,dQ = \mathcal{E}I\,dt$. Thus $\mathcal{E}I$ is the *rate* at which work is done on the circulating charges by whatever agency causes the nonelectrostatic force in the source. This term represents the rate of conversion of nonelectrical energy to electrical energy within the source. The term I^2r is the rate at which electrical energy is *dissipated* in the internal resistance of the source. The difference $\mathcal{E}I - I^2r$ is the *net* electrical power output of the source—that is, the rate at which the source delivers electrical energy to the remainder of the circuit.

Figure 25.22 Energy conversion in a simple circuit.

(a) Diagrammatic circuit



(b) A real circuit of the type shown in (a)



Power Input to a Source

Suppose that the lower rectangle in Fig. 25.22a is itself a source, with an emf *larger* than that of the upper source and opposite to that of the upper source. **Figure 25.23** shows a practical example, an automobile battery (the upper circuit element) being charged by the car's alternator (the lower element). The current I in the circuit is then *opposite* to that shown in Fig. 25.22; the lower source is pushing current backward through the upper source. Because of this reversal of current, instead of Eq. (25.15), we have for the upper source

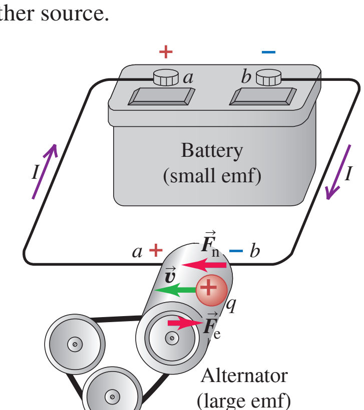
$$V_{ab} = \mathcal{E} + Ir$$

and instead of Eq. (25.19), we have

$$P = V_{ab}I = \mathcal{E}I + I^2r \tag{25.20}$$

Work is being done *on,* rather than *by,* the agent that causes the nonelectrostatic force in the upper source. There is a conversion of electrical energy into nonelectrical energy in the upper source at a rate $\mathcal{E}I$. The term I^2r in Eq. (25.20) is again the rate

Figure 25.23 When two sources are connected in a simple loop circuit, the source with the larger emf delivers energy to the other source.



of dissipation of energy in the internal resistance of the upper source, and the sum $\mathcal{E}I + I^2r$ is the total electrical power *input* to the upper source. This is what happens when a rechargeable battery (a storage battery) is connected to a charger. The charger supplies electrical energy to the battery; part of it is converted to chemical energy, to be reconverted later, and the remainder is dissipated (wasted) in the battery's internal resistance, warming the battery and causing a heat flow out of it. If you have a power tool or laptop computer with a rechargeable battery, you may have noticed that it gets warm while it is charging.

PROBLEM-SOLVING STRATEGY 25.1 Power and Energy in Circuits

IDENTIFY *the relevant concepts:* The ideas of electrical power input and output can be applied to any electric circuit. Many problems will ask you to explicitly consider power or energy.

SET UP *the problem* using the following steps:

1. Make a drawing of the circuit.
2. Identify the circuit elements, including sources of emf and resistors. We'll introduce other circuit elements later, including capacitors (Chapter 26) and inductors (Chapter 30).
3. Identify the target variables. Typically they will be the power input or output for each circuit element, or the total amount of energy put into or taken out of a circuit element in a given time.

EXECUTE *the solution* as follows:

1. A source of emf $\mathcal{E}$ delivers power $\mathcal{E}I$ into a circuit when current I flows through the source in the direction from $-$ to $+$. (For example, energy is converted from chemical energy in a battery, or from mechanical energy in a generator.) In this case there is a *positive* power output to the circuit or, equivalently, a *negative* power input to the source.
2. A source of emf takes power $\mathcal{E}I$ from a circuit when current passes through the source from $+$ to $-$. (This occurs in charging a storage battery, when electrical energy is converted to chemical energy.) In this case there is a *negative* power output to the circuit or, equivalently, a *positive* power input to the source.
3. There is always a *positive* power input to a resistor through which current flows, irrespective of the direction of current flow. This process removes energy from the circuit, converting it to heat at the rate $VI = I^2R = V^2/R$, where V is the potential difference across the resistor.
4. Just as in item 3, there always is a positive power input to the internal resistance r of a source through which current flows, irrespective of the direction of current flow. This process likewise removes energy from the circuit, converting it into heat at the rate I^2r.
5. If the power into or out of a circuit element is constant, the energy delivered to or extracted from that element is the product of power and elapsed time. (In Chapter 26 we'll encounter situations in which the power is not constant. In such cases, calculating the total energy requires an integral over the relevant time interval.)

EVALUATE *your answer:* Check your results; in particular, check that energy is conserved. This conservation can be expressed in either of two forms: "net power input = net power output" or "the algebraic sum of the power inputs to the circuit elements is zero."

EXAMPLE 25.8 Power input and output in a complete circuit

WITH VARIATION PROBLEMS

For the circuit that we analyzed in Example 25.5, find the rates of energy conversion (chemical to electrical) and energy dissipation in the battery, the rate of energy dissipation in the 4 Ω resistor, and the battery's net power output.

IDENTIFY and SET UP **Figure 25.24** shows the circuit, gives values of quantities known from Example 25.5, and indicates how we find the target variables. We use Eq. (25.19) to find the battery's net power output, the rate of chemical-to-electrical energy conversion, and the rate of energy dissipation in the battery's internal resistance. We use Eq. (25.18) to find the power delivered to (and dissipated in) the 4 Ω resistor.

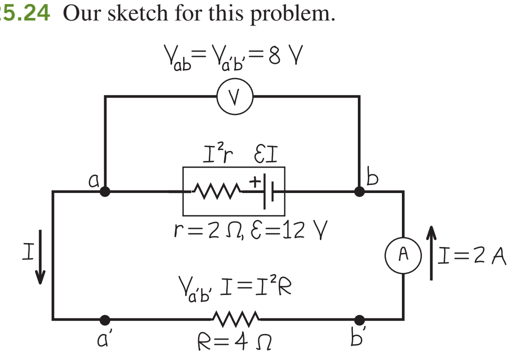
EXECUTE From the first term in Eq. (25.19), the rate of energy conversion in the battery is

$$\mathcal{E}I = (12\text{ V})(2\text{ A}) = 24\text{ W}$$

From the second term in Eq. (25.19), the rate of dissipation of energy in the battery is

$$I^2r = (2\text{ A})^2(2\ \Omega) = 8\text{ W}$$

Figure **25.24** Our sketch for this problem.



The *net* electrical power output of the battery is the difference between these: $\mathcal{E}I - I^2r = 16$ W. From Eq. (25.18), the electrical power input to, and the equal rate of dissipation of electrical energy in, the 4 Ω resistor are

$$V_{a'b'}I = (8\text{ V})(2\text{ A}) = 16\text{ W} \quad \text{and}$$
$$I^2R = (2\text{ A})^2(4\ \Omega) = 16\text{ W}$$

EVALUATE The rate $V_{a'b'}I$ at which energy is supplied to the 4 Ω resistor equals the rate I^2R at which energy is dissipated there. This is also equal to the battery's net power output: $P = V_{ab}I = (8\text{ V})(2\text{ A}) = 16\text{ W}$. In summary, the rate at which the source of emf supplies energy is $\mathcal{E}I = 24\text{ W}$, of which $I^2r = 8\text{ W}$ is dissipated in the battery's internal resistor and $I^2R = 16\text{ W}$ is dissipated in the external resistor.

KEYCONCEPT To find the electrical power input or output of any circuit element, multiply the current through the element by the potential difference across the element. There is a power output if the current flows through the element in the direction from low to high potential, and a power input if the current flows from high to low potential.

EXAMPLE 25.9 Increasing the resistance

WITH VARIATION PROBLEMS

Suppose we replace the external 4 Ω resistor in Fig. 25.24 with an 8 Ω resistor. How does this affect the electrical power dissipated in this resistor?

IDENTIFY and SET UP Our target variable is the power dissipated in the resistor to which the battery is connected. The situation is the same as in Example 25.8, but with a higher external resistance R.

EXECUTE According to Eq. (25.18), the power dissipated in the resistor is $P = I^2R$. You might conclude that making the resistance R twice as great as in Example 25.8 should make the power twice as great, or $2(16\text{ W}) = 32\text{ W}$. If instead you used the formula $P = V_{ab}{}^2/R$, you might conclude that the power should be one-half as great as in the preceding example, or $(16\text{ W})/2 = 8\text{ W}$. Which answer is correct?

In fact, *both* of these answers are *incorrect*. The first is wrong because changing the resistance R also changes the current in the circuit (remember, a source of emf does *not* generate the same current in all situations). The second answer is wrong because the potential difference V_{ab} across the resistor changes when the current changes. To get the correct answer, we first find the current just as we did in Example 25.5:

$$I = \frac{\mathcal{E}}{R + r} = \frac{12\text{ V}}{8\ \Omega + 2\ \Omega} = 1.2\text{ A}$$

The greater resistance causes the current to decrease. The potential difference across the resistor is

$$V_{ab} = IR = (1.2\text{ A})(8\ \Omega) = 9.6\text{ V}$$

which is greater than that with the 4 Ω resistor. We can then find the power dissipated in the resistor in either of two ways:

$$P = I^2R = (1.2\text{ A})^2(8\ \Omega) = 12\text{ W} \quad \text{or}$$

$$P = \frac{V_{ab}{}^2}{R} = \frac{(9.6\text{ V})^2}{8\ \Omega} = 12\text{ W}$$

EVALUATE Increasing the resistance R causes a *reduction* in the power input to the resistor. In the expression $P = I^2R$ the decrease in current is more important than the increase in resistance; in the expression $P = V_{ab}{}^2/R$ the increase in resistance is more important than the increase in V_{ab}. This same principle applies to ordinary light bulbs; a 50 W light bulb has a greater resistance than does a 100 W light bulb.

Can you show that replacing the 4 Ω resistor with an 8 Ω resistor decreases both the rate of energy conversion (chemical to electrical) in the battery and the rate of energy dissipation in the battery?

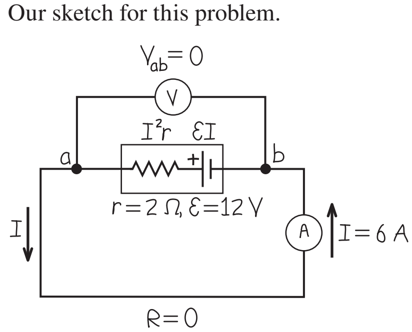
KEYCONCEPT The power input to a resistor (that is, the rate at which it dissipates energy) depends on its resistance R. For a given current through the resistor, increasing R increases the power; for a given potential difference across the resistor, increasing R decreases the power.

EXAMPLE 25.10 Power in a short circuit

For the short-circuit situation of Example 25.7, find the rates of energy conversion and energy dissipation in the battery and the net power output of the battery.

IDENTIFY and SET UP Our target variables are again the power inputs and outputs associated with the battery. **Figure 25.25** shows the circuit. This is the same situation as in Example 25.8, but now the external resistance R is zero.

Figure **25.25** Our sketch for this problem.



EXECUTE We found in Example 25.7 that the current in this situation is $I = 6\text{ A}$. From Eq. (25.19), the rate of energy conversion (chemical to electrical) in the battery is then

$$\mathcal{E}I = (12\text{ V})(6\text{ A}) = 72\text{ W}$$

and the rate of dissipation of energy in the battery is

$$I^2r = (6\text{ A})^2(2\ \Omega) = 72\text{ W}$$

The net power output of the source is $\mathcal{E}I - I^2r = 0$. We get this same result from the expression $P = V_{ab}I$, because the terminal voltage V_{ab} of the source is zero.

EVALUATE With ideal wires and an ideal ammeter, so that $R = 0$, *all* of the converted energy from the source is dissipated within the source. This is why a short-circuited battery is quickly ruined and may explode.

KEYCONCEPT Connecting the terminals of a source directly to each other is potentially unsafe because it can lead to a very high rate of energy dissipation in the source's internal resistance.

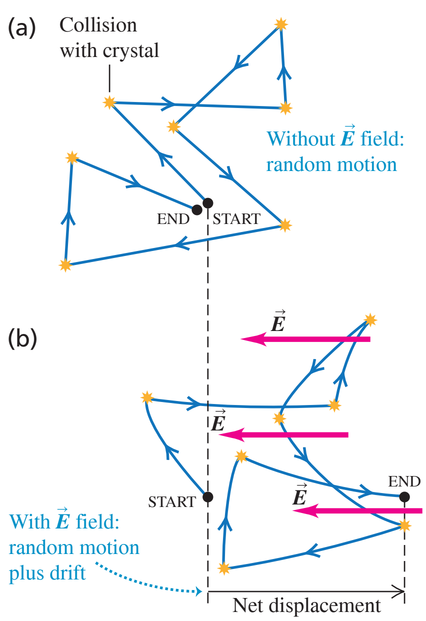
TEST YOUR UNDERSTANDING OF SECTION 25.5 Rank the following circuits in order from highest to lowest values of the net power output of the battery. (i) A 1.4 Ω resistor connected to a 1.5 V battery that has an internal resistance of 0.10 Ω; (ii) a 1.8 Ω resistor connected to a 4.0 V battery that has a terminal voltage of 3.6 V but an unknown internal resistance; (iii) an unknown resistor connected to a 12.0 V battery that has an internal resistance of 0.20 Ω and a terminal voltage of 11.0 V.

ANSWER

❙ **(iii), (ii), (i)** These are the same circuits that we analyzed in Test Your Understanding of Section 25.4. In each case the net power output of the battery is $P = V_{ab}I$, where V_{ab} is the battery terminal voltage. For circuit (i), we found that $I = 1.0\text{ A}$, so $V_{ab} = \mathcal{E} - Ir = 1.5\text{ V} - (1.0\text{ A})(0.10\ \Omega) = 1.4\text{ V}$, so $P = (1.4\text{ V})(1.0\text{ A}) = 1.4\text{ W}$. For circuit (ii), we have $V_{ab} = 3.6\text{ V}$ and found that $I = 2.0\text{ A}$, so $P = (3.6\text{ V})(2.0\text{ A}) = 7.2\text{ W}$. For circuit (iii), we have $V_{ab} = 11.0\text{ V}$ and found that $I = 5.0\text{ A}$, so $P = (11.0\text{ V})(5.0\text{ A}) = 55\text{ W}$.

25.6 THEORY OF METALLIC CONDUCTION

Figure **25.26** Random motions of an electron in a metallic crystal (a) with zero electric field and (b) with an electric field that causes drift. The curvatures of the paths are greatly exaggerated.



We can gain additional insight into electrical conduction by looking at the microscopic origin of conductivity. We'll consider a very simple model that treats the electrons as classical particles and ignores their quantum-mechanical behavior in solids. Using this model, we'll derive an expression for the resistivity of a metal. Even though this model is not entirely correct, it will still help you to develop an intuitive idea of the microscopic basis of conduction.

In the simplest microscopic model of conduction in a metal, each atom in the metallic crystal gives up one or more of its outer electrons. These electrons are then free to move through the crystal, colliding at intervals with the stationary positive ions. The motion of the electrons is analogous to the motion of molecules of a gas moving through a porous bed of sand.

If there is no electric field, the electrons move in straight lines between collisions, the directions of their velocities are random, and on average they never get anywhere (**Fig. 25.26a**). But if an electric field is present, the paths curve slightly because of the acceleration caused by electric-field forces. Figure 25.26b shows a few paths of an electron in an electric field directed from right to left. As we mentioned in Section 25.1, the average speed of random motion is of the order of 10^6 m/s, while the average drift speed is *much* slower, of the order of 10^{-4} m/s. The average time between collisions is called the **mean free time,** denoted by τ. **Figure 25.27** shows a mechanical analog of this electron motion.

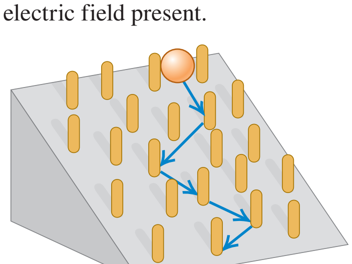
We would like to derive from this model an expression for the resistivity ρ of a material, defined by Eq. (25.5):

$$\rho = \frac{E}{J} \tag{25.21}$$

where E and J are the magnitudes of electric field and current density, respectively. The current density $\vec{J}$ is in turn given by Eq. (25.4):

$$\vec{J} = nq\vec{v}_d \tag{25.22}$$

Figure **25.27** The motion of a ball rolling down an inclined plane and bouncing off pegs in its path is analogous to the motion of an electron in a metallic conductor with an electric field present.



where n is the number of free electrons per unit volume (the electron concentration), $q = -e$ is the charge of each, and $\vec{v}_d$ is their average drift velocity.

We need to relate the drift velocity $\vec{v}_d$ to the electric field $\vec{E}$. The value of $\vec{v}_d$ is determined by a steady-state condition in which, on average, the velocity *gains* of the charges due to the force of the $\vec{E}$ field are just balanced by the velocity *losses* due to collisions. To clarify this process, let's imagine turning on the two effects one at a time. Suppose that before time $t = 0$ there is no field. The electron motion is then completely random. A typical electron has velocity $\vec{v}_0$ at time $t = 0$, and the value of $\vec{v}_0$ averaged over many electrons (that is, the initial velocity of an average electron) is zero: $(\vec{v}_0)_{av} = \mathbf{0}$. Then at time $t = 0$ we turn on a constant electric field $\vec{E}$. The field exerts a force $\vec{F} = q\vec{E}$ on each charge, and this causes an acceleration $\vec{a}$ in the direction of the force, given by

$$\vec{a} = \frac{\vec{F}}{m} = \frac{q\vec{E}}{m}$$

where m is the electron mass. *Every* electron has this acceleration.

After a time τ, the average time between collisions, we "turn on" the collisions. At time $t = \tau$ an electron that has velocity $\vec{v}_0$ at time $t = 0$ has a velocity

$$\vec{v} = \vec{v}_0 + \vec{a}\tau$$

The velocity $\vec{v}_{av}$ of an *average* electron at this time is the sum of the averages of the two terms on the right. As we have pointed out, the initial velocity $\vec{v}_0$ is zero for an average electron, so

$$\vec{v}_{av} = \vec{a}\tau = \frac{q\tau}{m}\vec{E} \tag{25.23}$$

After time $t = \tau$, the tendency of the collisions to decrease the velocity of an average electron (by means of randomizing collisions) just balances the tendency of the $\vec{E}$ field to increase this velocity. Thus the velocity of an average electron, given by Eq. (25.23), is maintained over time and is equal to the drift velocity $\vec{v}_d$:

$$\vec{v}_d = \frac{q\tau}{m}\vec{E}$$

Now we substitute this equation for the drift velocity $\vec{v}_d$ into Eq. (25.22):

$$\vec{J} = nq\vec{v}_d = \frac{nq^2\tau}{m}\vec{E}$$

Comparing this with Eq. (25.21), which we can rewrite as $\vec{J} = \vec{E}/\rho$, and substituting $q = -e$ for an electron, we see that

Resistivity of a metal

$$\rho = \frac{m}{ne^2\tau} \tag{25.24}$$

Electron mass

Average time between collisions

Number of free electrons per unit volume

Magnitude of electron charge

If n and τ are independent of $\vec{E}$, then the resistivity is independent of $\vec{E}$ and the conducting material obeys Ohm's law.

Turning the interactions on one at a time may seem artificial. But the derivation would come out the same if each electron had its own clock and the $t = 0$ times were different for different electrons. If τ is the average time between collisions, then $\vec{v}_d$ is still the average electron drift velocity, even though the motions of the various electrons aren't actually correlated in the way we postulated.

What about the temperature dependence of resistivity? In a perfect crystal with no atoms out of place, a correct quantum-mechanical analysis would let the free electrons move through the crystal with no collisions at all. But the atoms vibrate about their equilibrium positions. As the temperature increases, the amplitudes of these vibrations increase, collisions become more frequent, and the mean free time τ decreases. So this theory predicts that the resistivity of a metal increases with temperature. In a superconductor, roughly speaking, there are no inelastic collisions, τ is infinite, and the resistivity ρ is zero.

In a pure semiconductor such as silicon or germanium, the number of charge carriers per unit volume, n, is not constant but increases very rapidly with increasing temperature. This increase in n far outweighs the decrease in the mean free time, and in a semiconductor the resistivity always decreases rapidly with increasing temperature. At low temperatures, n is very small, and the resistivity becomes so large that the material can be considered an insulator.

Electrons gain energy between collisions through the work done on them by the electric field. During collisions they transfer some of this energy to the atoms of the material of the conductor. This leads to an increase in the material's internal energy and temperature; that's why wires carrying current get warm. If the electric field in the material is large enough, an electron can gain enough energy between collisions to knock off electrons that are normally bound to atoms in the material. These can then knock off more electrons, and so on, leading to an avalanche of current. This is the basis of dielectric breakdown in insulators (see Section 24.4).

EXAMPLE 25.11 Mean free time in copper

Calculate the mean free time between collisions in copper at room temperature.

IDENTIFY and SET UP We can obtain an expression for mean free time τ in terms of n, ρ, e, and m by rearranging Eq. (25.24). From Example 25.1 and Table 25.1, for copper $n = 8.5 \times 10^{28}\ \text{m}^{-3}$ and $\rho = 1.72 \times 10^{-8}\ \Omega \cdot \text{m}$. In addition, $e = 1.60 \times 10^{-19}\ \text{C}$ and $m = 9.11 \times 10^{-31}\ \text{kg}$ for electrons.

EXECUTE From Eq. (25.24), we get

$$\begin{aligned}\tau &= \frac{m}{ne^2\rho}\\ &= \frac{9.11 \times 10^{-31}\ \text{kg}}{(8.5 \times 10^{28}\ \text{m}^{-3})(1.60 \times 10^{-19}\ \text{C})^2(1.72 \times 10^{-8}\ \Omega \cdot \text{m})}\\ &= 2.4 \times 10^{-14}\ \text{s}\end{aligned}$$

EVALUATE The mean free time is the average time between collisions for a given electron. Taking the reciprocal, we find that each electron averages $1/\tau = 4.2 \times 10^{13}$ collisions per second!

KEYCONCEPT The resistivity of a metal is related to the average time that a current-carrying electron can travel through the metal without colliding with one of the metal's ions. The greater this mean free time, the lower the resistivity.

TEST YOUR UNDERSTANDING OF SECTION 25.6 Which of the following factors will, if increased, make it more difficult to produce a certain amount of current in a conductor? (There may be more than one correct answer.) (i) The mass of the moving charged particles in the conductor; (ii) the number of moving charged particles per cubic meter; (iii) the amount of charge on each moving particle; (iv) the average time between collisions for a typical moving charged particle.

ANSWER

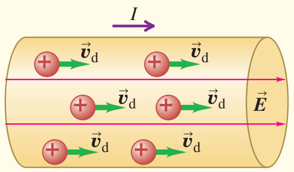
❙ (i) The difficulty of producing a certain amount of current increases as the resistivity ρ increases. From Eq. (25.24), $\rho = m/ne^2\tau$, so increasing the mass m will increase the resistivity. That's because a more massive charged particle will respond more sluggishly to an applied electric field and hence drift more slowly. To produce the same current, a greater electric field would be needed. (Increasing n, e, or τ would decrease the resistivity and make it easier to produce a given current.)

CHAPTER 25 SUMMARY

Current and current density: Current is the amount of charge flowing through a specified area, per unit time. The SI unit of current is the ampere $(1\text{ A} = 1\text{ C/s})$. The current I through an area A depends on the concentration n and charge q of the charge carriers, as well as on the magnitude of their drift velocity $\vec{v}_d$. The current density is current per unit cross-sectional area. Current is usually described in terms of a flow of positive charge, even when the charges are actually negative or of both signs. (See Example 25.1.)

$$I = \frac{dQ}{dt} = n|q|v_d A \quad (25.2)$$

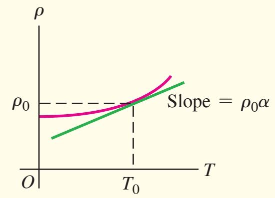
$$\vec{J} = nq\vec{v}_d \quad (25.4)$$



Resistivity: The resistivity ρ of a material is the ratio of the magnitudes of electric field and current density. Good conductors have small resistivity; good insulators have large resistivity. Ohm's law, obeyed approximately by many materials, states that ρ is a constant independent of the value of E. Resistivity usually increases with temperature; for small temperature changes this variation is represented approximately by Eq. (25.6), where α is the temperature coefficient of resistivity.

$$\rho = \frac{E}{J} \quad (25.5)$$

$$\rho(T) = \rho_0[1 + \alpha(T - T_0)] \quad (25.6)$$

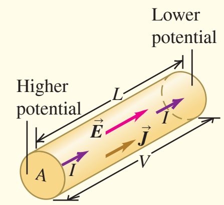


Metal: ρ increases with increasing T.

Resistors: The potential difference V across a sample of material that obeys Ohm's law is proportional to the current I through the sample. The ratio $V/I = R$ is the resistance of the sample. The SI unit of resistance is the ohm $(1\ \Omega = 1\text{ V/A})$. The resistance of a cylindrical conductor is related to its resistivity ρ, length L, and cross-sectional area A. (See Examples 25.2 and 25.3.)

$$V = IR \quad (25.11)$$

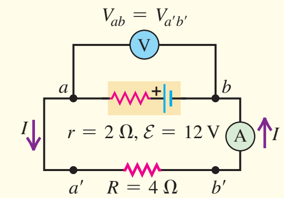
$$R = \frac{\rho L}{A} \quad (25.10)$$



Circuits and emf: A complete circuit has a continuous current-carrying path. A complete circuit carrying a steady current must contain a source of electromotive force (emf) $\mathcal{E}$. The SI unit of electromotive force is the volt (V). Every real source of emf has some internal resistance r, so its terminal potential difference V_{ab} depends on current. (See Examples 25.4–25.7.)

$$V_{ab} = \mathcal{E} - Ir \quad (25.15)$$

(source with internal resistance)



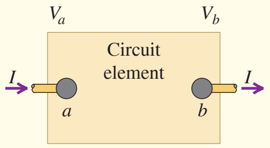
Energy and power in circuits: A circuit element puts energy into a circuit if the current direction is from lower to higher potential in the device, and it takes energy out of the circuit if the current is opposite. The power P equals the product of the potential difference $V_a - V_b = V_{ab}$ and the current I. A resistor always takes electrical energy out of a circuit. (See Examples 25.8–25.10.)

$$P = V_{ab}I \quad (25.17)$$

(general circuit element)

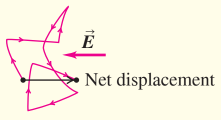
$$P = V_{ab}I = I^2R = \frac{V_{ab}^2}{R} \quad (25.18)$$

(power delivered to a resistor)



Conduction in metals: In a metal, current is due to the motion of electrons. They move freely through the metallic crystal but collide with positive ions. In a crude classical model of this motion, the resistivity of the material can be related to the electron mass, charge, speed of random motion, density, and mean free time between collisions. (See Example 25.11.)

$$\rho = \frac{m}{ne^2\tau} \quad (25.24)$$



Chapter 25 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 25.2 and 25.3 (Section 25.3) before attempting these problems.

VP25.3.1 Some airliners use wires of aluminum (resistivity $2.75 \times 10^{-8}\ \Omega \cdot \text{m}$) rather than copper to reduce weight. The potential difference between the ends of an aluminum wire 35.0 m long is 1.45 V when the wire carries current 1.20 A. Find (a) the electric-field magnitude inside the wire, (b) the resistance of the wire, and (c) the cross-sectional area of the wire.

VP25.3.2 When a current of 0.218 A is present in a wire 20.0 m in length with cross-sectional area $8.19 \times 10^{-7}\ \text{m}^2$, the electric field inside the wire is 6.50×10^{-3} V/m. What are (a) the resistance of the wire and (b) the resistivity of the material of which the wire is made?

VP25.3.3 Two wires of circular cross section are made of the same metal. Wire 1 has radius r and length L; wire 2 has radius $2r$ and length $2L$. The potential difference between the ends of the wire is the same for both wires. For each of the following quantities, find the ratio of its value for wire 2 to its value for wire 1: (a) resistance of the wire; (b) current in the wire; (c) current density in the wire; (d) electric-field magnitude in the wire.

VP25.3.4 The current in a certain wire at 20.0°C is 2.40 A when the potential difference between the ends of the wire is 1.50 V. If the wire is heated to 80.0°C and the potential difference remains the same, the current decreases to 2.00 A. (a) What is the resistance of the wire at 20.0°C? At 80.0°C? (b) What is the temperature coefficient of resistivity for the material of which this wire is made?

Be sure to review EXAMPLE 25.5 (Section 25.4) before attempting these problems.

VP25.5.1 A battery with emf 9.00 V and internal resistance 1.10 Ω is in a complete circuit with a resistor of resistance 15.3 Ω. Find (a) the current in the circuit, (b) the potential difference across the resistor, and (c) the potential difference across the battery terminals.

VP25.5.2 When an automobile battery with an emf of 12.6 V is connected to a resistor of resistance 25.0 Ω, the current in the circuit is 0.480 A. Find (a) the potential difference across the resistor and (b) the internal resistance of the battery.

VP25.5.3 An AA battery has an emf of 1.50 V. After being used for quite a while, one AA battery has a potential difference across its terminals of 1.30 V when it is connected to a light bulb with resistance 20.0 Ω. Find (a) the current in this circuit and (b) the internal resistance of the battery.

VP25.5.4 The manufacturer of one brand of AAA battery states that the battery's emf is 1.50 V, its internal resistance is 0.160 Ω, and the maximum allowed current for short times is 2.00 A. Find (a) the smallest resistance to which this battery can be connected for short times and (b) the potential difference across the battery terminals when it is so connected.

Be sure to review EXAMPLES 25.8 and 25.9 (Section 25.5) before attempting these problems.

VP25.9.1 A battery with emf 24.0 V and internal resistance 1.30 Ω is in a complete circuit with a resistor of resistance 18.0 Ω. Find (a) the current in the circuit, (b) the rate at which energy is converted from chemical to electrical form in the battery, (c) the rate of energy dissipation in the battery, (d) the net power output of the battery, and (e) the rate at which energy is dissipated in the 18.0 Ω resistor.

VP25.9.2 When you connect a battery with emf 9.00 V to a resistor of resistance 12.0 Ω, energy is dissipated in the resistor at the rate of 6.54 W. Find (a) the current in the circuit, (b) the net power output of the battery, (c) the potential difference across the terminals of the battery, and (d) the internal resistance of the battery.

VP25.9.3 The current is 0.720 A in a circuit in which a battery of emf 12.0 V is connected to a resistor of resistance 16.0 Ω. Find (a) the potential difference across the terminals of the battery, (b) the internal resistance of the battery, (c) the power dissipated in the resistor, (d) the rate at which chemical energy is converted to electrical energy in the battery, and (e) the rate of energy dissipation in the battery.

VP25.9.4 A battery with emf $\mathcal{E}$ and internal resistance r is first connected to a resistor of resistance $R_1 = R$, where R is greater than r. This resistor is then disconnected from the battery and replaced with a resistor of resistance $R_2 = 2R$. Find the ratio of (a) the current with the second resistor to the current with the first resistor and (b) the power dissipated in the second resistor to the power dissipated in the first resistor. (c) In which resistor is the current greater? In which is the power dissipated greater?

BRIDGING PROBLEM Resistivity, Temperature, and Power

A toaster using a Nichrome heating element operates on 120 V. When it is switched on at 20°C, the heating element carries an initial current of 1.35 A. A few seconds later the current reaches the steady value of 1.23 A. (a) What is the final temperature of the element? The average value of the temperature coefficient of resistivity for Nichrome over the relevant temperature range is $4.5 \times 10^{-4}\ (\text{C°})^{-1}$. (b) What is the power dissipated in the heating element initially and when the current reaches 1.23 A?

SOLUTION GUIDE

IDENTIFY and SET UP

1. A heating element acts as a resistor that converts electrical energy into thermal energy. The resistivity ρ of Nichrome depends on temperature, and hence so does the resistance $R = \rho L/A$ of the heating element and the current $I = V/R$ that it carries.
2. We are given $V = 120$ V and the initial and final values of I. Select an equation that will allow you to find the initial and final values of resistance, and an equation that relates resistance to temperature [the target variable in part (a)].
3. The power P dissipated in the heating element depends on I and V. Select an equation that will allow you to calculate the initial and final values of P.

EXECUTE

4. Combine your equations from step 2 to give a relationship between the initial and final values of I and the initial and final temperatures (20°C and T_{final}).
5. Solve your expression from step 4 for T_{final}.
6. Use your equation from step 3 to find the initial and final powers.

EVALUATE

7. Is the final temperature greater than or less than 20°C? Does this make sense?
8. Is the final resistance greater than or less than the initial resistance? Again, does this make sense?
9. Is the final power greater than or less than the initial power? Does this agree with your observations in step 8?

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q25.1 The definition of resistivity $(\rho = E/J)$ implies that an electric field exists inside a conductor. Yet we saw in Chapter 21 that there can be no electrostatic electric field inside a conductor. Is there a contradiction here? Explain.

Q25.2 A cylindrical rod has resistance R. If we triple its length and diameter, what is its resistance, in terms of R?

Q25.3 A cylindrical rod has resistivity ρ. If we triple its length and diameter, what is its resistivity, in terms of ρ?

Q25.4 Two copper wires with different diameters are joined end to end. If a current flows in the wire combination, what happens to electrons when they move from the larger-diameter wire into the smaller-diameter wire? Does their drift speed increase, decrease, or stay the same? If the drift speed changes, what is the force that causes the change? Explain your reasoning.

Q25.5 When is a 1.5 V AAA battery *not* actually a 1.5 V battery? That is, when do its terminals provide a potential difference of less than 1.5 V?

Q25.6 Can the potential difference between the terminals of a battery ever be opposite in direction to the emf? If it can, give an example. If it cannot, explain why not.

Q25.7 A rule of thumb used to determine the internal resistance of a source is that it is the open-circuit voltage divided by the short-circuit current. Is this correct? Why or why not?

Q25.8 Batteries are always labeled with their emf; for instance, an AA flashlight battery is labeled "1.5 volts." Would it also be appropriate to put a label on batteries stating how much current they provide? Why or why not?

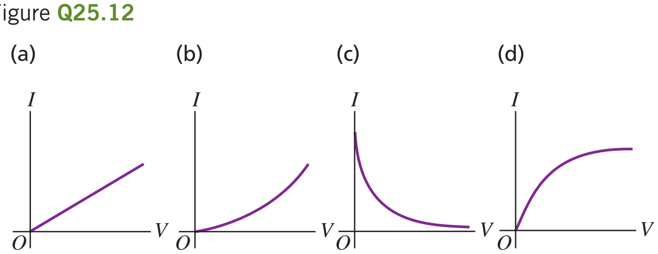
Q25.9 We have seen that a coulomb is an enormous amount of charge; it is virtually impossible to place a charge of 1 C on an object. Yet, a current of 10 A, 10 C/s, is quite reasonable. Explain this apparent discrepancy.

Q25.10 Electrons in an electric circuit pass through a resistor. The wire on either side of the resistor has the same diameter. (a) How does the drift speed of the electrons before entering the resistor compare to the speed after leaving the resistor? Explain your reasoning. (b) How does the potential energy for an electron before entering the resistor compare to the potential energy after leaving the resistor? Explain your reasoning.

Q25.11 Temperature coefficients of resistivity are given in Table 25.2. (a) If a copper heating element is connected to a source of constant voltage, does the electrical power consumed by the heating element increase or decrease as its temperature increases? Explain. (b) A resistor in the form of a carbon cylinder is connected to the voltage source. As the temperature of the cylinder increases, does the electrical power it consumes increase or decrease? Explain.

Q25.12 Which of the graphs in **Fig. Q25.12** best illustrates the current I in a real resistor as a function of the potential difference V across it? Explain.

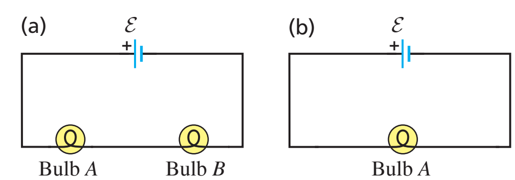
Figure **Q25.12**



Q25.13 Why does an electric light bulb nearly always burn out just as you turn on the light, almost never while the light is shining?

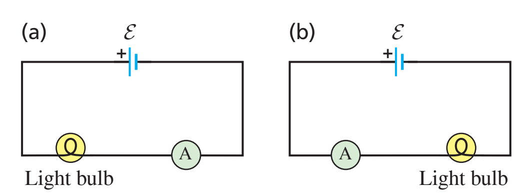
Q25.14 A light bulb glows because it has resistance. The brightness of a light bulb increases with the electrical power dissipated in the bulb. (a) In the circuit shown in **Fig. Q25.14a**, the two bulbs A and B are identical. Compared to bulb A, does bulb B glow more brightly, just as brightly, or less brightly? Explain your reasoning. (b) Bulb B is removed from the circuit and the circuit is completed as shown in Fig. Q25.14b. Compared to the brightness of bulb A in Fig. Q25.14a, does bulb A now glow more brightly, just as brightly, or less brightly? Explain your reasoning.

Figure **Q25.14**



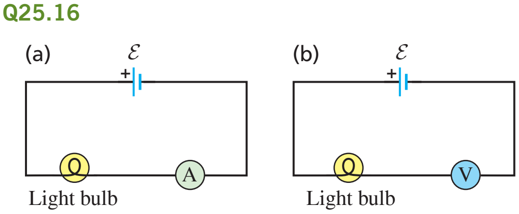
Q25.15 (See Discussion Question Q25.14.) An ideal ammeter A is placed in a circuit with a battery and a light bulb as shown in **Fig. Q25.15a**, and the ammeter reading is noted. The circuit is then reconnected as in Fig. Q25.15b, so that the positions of the ammeter and light bulb are reversed. (a) How does the ammeter reading in the situation shown in Fig. Q25.15a compare to the reading in the situation shown in Fig. Q25.15b? Explain your reasoning. (b) In which situation does the light bulb glow more brightly? Explain your reasoning.

Figure **Q25.15**



Q25.16 (See Discussion Question Q25.14.) Will a light bulb glow more brightly when it is connected to a battery as shown in **Fig. Q25.16a**, in which an ideal ammeter A is placed in the circuit, or when it is connected as shown in Fig. 25.16b, in which an ideal voltmeter V is placed in the circuit? Explain your reasoning.

Figure **Q25.16**



Q25.17 The energy that can be extracted from a storage battery is always less than the energy that goes into it while it is being charged. Why?

Q25.18 Eight flashlight batteries in series have an emf of about 12 V, similar to that of a car battery. Could they be used to start a car with a dead battery? Why or why not?

Q25.19 Small aircraft often have 24 V electrical systems rather than the 12 V systems in automobiles, even though the electrical power requirements are roughly the same in both applications. The explanation given by aircraft designers is that a 24 V system weighs less than a 12 V system because thinner wires can be used. Explain why this is so.

Q25.20 Long-distance, electric-power, transmission lines always operate at very high voltage, sometimes as much as 750 kV. What are the advantages of such high voltages? What are the disadvantages?

Q25.21 Ordinary household electric lines in North America usually operate at 120 V. Why is this a desirable voltage, rather than a value considerably larger or smaller? On the other hand, automobiles usually have 12 V electrical systems. Why is this a desirable voltage?

Q25.22 A fuse is a device designed to break a circuit, usually by melting when the current exceeds a certain value. What characteristics should the material of the fuse have?

Q25.23 High-voltage power supplies are sometimes designed intentionally to have rather large internal resistance as a safety precaution. Why is such a power supply with a large internal resistance safer than a supply with the same voltage but lower internal resistance?

Q25.24 The text states that good thermal conductors are also good electrical conductors. If so, why don't the cords used to connect toasters, irons, and similar heat-producing appliances get hot by conduction of heat from the heating element?

EXERCISES

Section 25.1 Current

25.1 • **Lightning Strikes.** During lightning strikes from a cloud to the ground, currents as high as 25,000 A can occur and last for about 40 μs. How much charge is transferred from the cloud to the earth during such a strike?

25.2 • A silver wire 2.6 mm in diameter transfers a charge of 420 C in 80 min. Silver contains 5.8×10^{28} free electrons per cubic meter. (a) What is the current in the wire? (b) What is the magnitude of the drift velocity of the electrons in the wire?

25.3 • A 5.00 A current runs through a 12 gauge copper wire (diameter 2.05 mm) and through a light bulb. Copper has 8.5×10^{28} free electrons per cubic meter. (a) How many electrons pass through the light bulb each second? (b) What is the current density in the wire? (c) At what speed does a typical electron pass by any given point in the wire? (d) If you were to use wire of twice the diameter, which of the above answers would change? Would they increase or decrease?

25.4 • An 18 gauge copper wire (diameter 1.02 mm) carries a current with a current density of $3.20 \times 10^6\ \mathrm{A/m^2}$. The density of free electrons for copper is 8.5×10^{28} electrons per cubic meter. Calculate (a) the current in the wire and (b) the magnitude of the drift velocity of electrons in the wire.

25.5 • The free-electron density in a copper wire is $8.5 \times 10^{28}\ \mathrm{electrons/m^3}$. The electric field in the wire is 0.0600 N/C and the temperature of the wire is 20.0°C. (a) What is the drift speed v_d of the electrons in the wire? (b) What is the potential difference between two points in the wire that are separated by 20.0 cm?

25.6 •• You want to produce three 1.00-mm-diameter cylindrical wires, each with a resistance of 1.00 Ω at room temperature. One wire is gold, one is copper, and one is aluminum. Refer to Table 25.1 for the resistivity values. (a) What will be the length of each wire? (b) Gold has a density of $1.93 \times 10^4\ \mathrm{kg/m^3}$. What will be the mass of the gold wire? If you consider the current price of gold, is this wire very expensive?

25.7 • **CALC** The current in a wire varies with time according to the relationship $I = 55\ \mathrm{A} - (0.65\ \mathrm{A/s^2})t^2$. (a) How many coulombs of charge pass a cross section of the wire in the time interval between $t = 0$ and $t = 8.0$ s? (b) What constant current would transport the same charge in the same time interval?

25.8 • Current passes through a solution of sodium chloride. In 1.00 s, 2.68×10^{16} Na^+ ions arrive at the negative electrode and 3.92×10^{16} Cl^- ions arrive at the positive electrode. (a) What is the current passing between the electrodes? (b) What is the direction of the current?

25.9 • **BIO** **Transmission of Nerve Impulses.** Nerve cells transmit electric signals through their long tubular axons. These signals propagate due to a sudden rush of Na^+ ions, each with charge $+e$, into the axon. Measurements have revealed that typically about 5.6×10^{11} Na^+ ions enter each meter of the axon during a time of 10 ms. What is the current during this inflow of charge in a meter of axon?

25.10 •• The mass density of silver at room temperature is $10.5 \times 10^3\ \mathrm{kg/m^3}$ and its atomic mass is 108 g/mol. If we assume there is one free electron per silver atom, what is the free-electron density for silver, in $\mathrm{electrons/m^3}$? How does your answer compare to the free-electron density for copper given in Example 25.1?

Section 25.2 Resistivity and Section 25.3 Resistance

25.11 •• A metal wire has a circular cross section with radius 0.800 mm. You measure the resistivity of the wire in the following way: You connect one end of the wire to one terminal of a battery that has emf 12.0 V and negligible internal resistance. To the other terminal of the battery you connect a point along the wire so that the length of wire between the battery terminals is d. You measure the current in the wire as a function of d. The currents are small, so the temperature change of the wire is very small. You plot your results as I versus $1/d$ and find that the data lie close to a straight line that has slope 600 A · m. What is the resistivity of the material of which the wire is made?

25.12 • (a) At room temperature, what is the strength of the electric field in a 12 gauge copper wire (diameter 2.05 mm) that is needed to cause a 4.50 A current to flow? (b) What field would be needed if the wire were made of silver instead?

25.13 •• A 1.50 m cylindrical rod of diameter 0.500 cm is connected to a power supply that maintains a constant potential difference of 15.0 V across its ends, while an ammeter measures the current through it. You observe that at room temperature (20.0°C) the ammeter reads 18.5 A, while at 92.0°C it reads 17.2 A. You can ignore any thermal expansion of the rod. Find (a) the resistivity at 20.0°C and (b) the temperature coefficient of resistivity at 20°C for the material of the rod.

25.14 • A copper wire has a square cross 2.3 mm on a side. The wire is 4.0 m long and carries a current of 3.6 A. The density of free electrons is $8.5 \times 10^{28}/\mathrm{m^3}$. Find the magnitudes of (a) the current density in the wire and (b) the electric field in the wire. (c) How much time is required for an electron to travel the length of the wire?

25.15 • A 14 gauge copper wire of diameter 1.628 mm carries a current of 12.5 mA. (a) What is the potential difference across a 2.00 m length of the wire? (b) What would the potential difference in part (a) be if the wire were silver instead of copper, but all else were the same?

25.16 •• A wire 6.50 m long with diameter of 2.05 mm has a resistance of 0.0290 Ω. What material is the wire most likely made of?

25.17 • A copper wire has radius 0.800 mm and carries current I at 20.0°C. A silver wire with radius 0.500 mm carries the same current and is also at 20.0°C. Let E_{Cu} be the electric field in the copper wire and E_{Ag} be the electric field in the silver wire. What is the ratio E_{Cu}/E_{Ag}?

25.18 •• A ductile metal wire has resistance R. What will be the resistance of this wire in terms of R if it is stretched to three times its original length, assuming that the density and resistivity of the material do not change when the wire is stretched? (*Hint:* The amount of metal does not change, so stretching out the wire will affect its cross-sectional area.)

25.19 • In household wiring, copper wire 2.05 mm in diameter is often used. Find the resistance of a 24.0 m length of this wire.

25.20 •• What diameter must a copper wire have if its resistance is to be the same as that of an equal length of aluminum wire with diameter 2.14 mm?

25.21 • A current-carrying gold wire has diameter 0.84 mm. The electric field in the wire is 0.49 V/m. What are (a) the current carried by the wire; (b) the potential difference between two points in the wire 6.4 m apart; (c) the resistance of a 6.4 m length of this wire?

25.22 • You apply a potential difference of 4.50 V between the ends of a wire that is 2.50 m in length and 0.654 mm in radius. The resulting current through the wire is 17.6 A. What is the resistivity of the wire?

25.23 • (a) What is the resistance of a Nichrome wire at 0.0°C if its resistance is 100.00 Ω at 11.5°C? (b) What is the resistance of a carbon rod at 25.8°C if its resistance is 0.0160 Ω at 0.0°C?

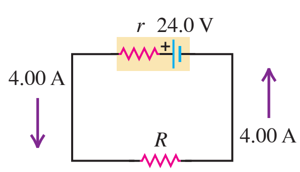
25.24 • A carbon resistor is to be used as a thermometer. On a winter day when the temperature is 4.0°C, the resistance of the carbon resistor is 217.3 Ω. What is the temperature on a spring day when the resistance is 215.8 Ω? (Take the reference temperature T_0 to be 4.0°C.)

Section 25.4 Electromotive Force and Circuits

25.25 • A copper transmission cable 100 km long and 10.0 cm in diameter carries a current of 125 A. (a) What is the potential drop across the cable? (b) How much electrical energy is dissipated as thermal energy every hour?

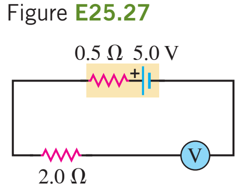
25.26 • Consider the circuit shown in **Fig. E25.26**. The terminal voltage of the 24.0 V battery is 21.2 V. What are (a) the internal resistance r of the battery and (b) the resistance R of the circuit resistor?

Figure **E25.26**



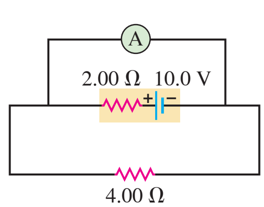
25.27 • An ideal voltmeter V is connected to a 2.0 Ω resistor and a battery with emf 5.0 V and internal resistance 0.5 Ω as shown in **Fig. E25.27**. (a) What is the current in the 2.0 Ω resistor? (b) What is the terminal voltage of the battery? (c) What is the reading on the voltmeter? Explain your answers.

Figure **E25.27**



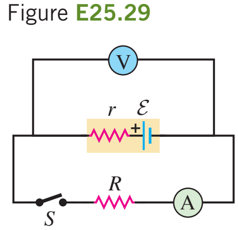
25.28 • An idealized ammeter is connected to a battery as shown in **Fig. E25.28**. Find (a) the reading of the ammeter, (b) the current through the 4.00 Ω resistor, (c) the terminal voltage of the battery.

Figure **E25.28**



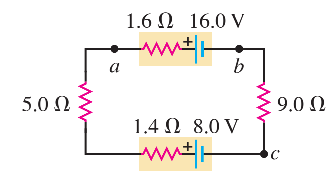
25.29 • When switch S in **Fig. E25.29** is open, the voltmeter V reads 3.08 V. When the switch is closed, the voltmeter reading drops to 2.97 V, and the ammeter A reads 1.65 A. Find the emf, the internal resistance of the battery, and the circuit resistance R. Assume that the two meters are ideal, so they don't affect the circuit.

Figure **E25.29**



25.30 • The circuit shown in **Fig. E25.30** contains two batteries, each with an emf and an internal resistance, and two resistors. Find (a) the current in the circuit (magnitude *and* direction); (b) the terminal voltage V_{ab} of the 16.0 V battery; (c) the potential difference V_{ac} of point a with respect to point c. (d) Using Fig. 25.20 as a model, graph the potential rises and drops in this circuit.

Figure **E25.30**



25.31 •• In the circuit shown in **Fig. E25.30**, the 16.0 V battery is removed and reinserted with the opposite polarity, so its negative terminal is now next to point a. Find (a) the current in the circuit (magnitude and direction) and (b) the terminal voltage V_{ab} of the 16.0 V battery.

25.32 • A battery has emf 30.0 V and internal resistance r. A 9.00 Ω resistor is connected to the terminals of the battery, and the voltage drop across the resistor is 27.0 V. What is the internal resistance of the battery?

Section 25.5 Energy and Power in Electric Circuits

25.33 •• A battery has emf 24.0 V and internal resistance 3.00 Ω. A resistor of resistance R is connected to the battery. What are the two values of R for which 21.0 W of electrical power is consumed in the resistor?

25.34 •• When a resistor with resistance R is connected to a 1.50 V flashlight battery, the resistor consumes 0.0625 W of electrical power. (Throughout, assume that each battery has negligible internal resistance.) (a) What power does the resistor consume if it is connected to a 12.6 V car battery? Assume that R remains constant when the power consumption changes. (b) The resistor is connected to a battery and consumes 5.00 W. What is the voltage of this battery?

25.35 • **Light Bulbs.** The power rating of a light bulb (such as a 100 W bulb) is the power it dissipates when connected across a 120 V potential difference. What is the resistance of (a) a 100 W bulb and (b) a 60 W bulb? (c) How much current does each bulb draw in normal use?

25.36 • If a "75 W" bulb (see Problem 25.35) is connected across a 220 V potential difference (as is used in Europe), how much power does it dissipate? Ignore the temperature dependence of the bulb's resistance.

25.37 • **European Light Bulb.** In Europe the standard voltage in homes is 220 V instead of the 120 V used in the United States. Therefore a "100 W" European bulb would be intended for use with a 220 V potential difference (see Problem 25.36). (a) If you bring a "100 W" European bulb home to the United States, what should be its U.S. power rating? (b) How much current will the 100 W European bulb draw in normal use in the United States?

25.38 • A battery-powered global positioning system (GPS) receiver operating on 9.0 V draws a current of 0.13 A. How much electrical energy does it consume during 30 minutes?

25.39 • Consider the circuit of Fig. E25.30. (a) What is the total rate at which electrical energy is dissipated in the 5.0 Ω and 9.0 Ω resistors? (b) What is the power output of the 16.0 V battery? (c) At what rate is electrical energy being converted to other forms in the 8.0 V battery? (d) Show that the power output of the 16.0 V battery equals the overall rate of consumption of electrical energy in the rest of the circuit.

25.40 • BIO **Electric Eels.** Electric eels generate electric pulses along their skin that can be used to stun an enemy when they come into contact with it. Tests have shown that these pulses can be up to 500 V and produce currents of 80 mA (or even larger). A typical pulse lasts for 10 ms. What power and how much energy are delivered to the unfortunate enemy with a single pulse, assuming a steady current?

25.41 • BIO **Treatment of Heart Failure.** A heart defibrillator is used to enable the heart to start beating if it has stopped. This is done by passing a large current of 12 A through the body at 25 V for a very short time, usually about 3.0 ms. (a) What power does the defibrillator deliver to the body, and (b) how much energy is transferred?

25.42 •• The battery for a certain cell phone is rated at 3.70 V. According to the manufacturer it can produce 3.15×10^4 J of electrical energy, enough for 5.25 h of operation, before needing to be recharged. Find the average current that this cell phone draws when turned on.

25.43 •• The capacity of a storage battery, such as those used in automobile electrical systems, is rated in ampere-hours $(\text{A} \cdot \text{h})$. A 50 $\text{A} \cdot \text{h}$ battery can supply a current of 50 A for 1.0 h, or 25 A for 2.0 h, and so on. (a) What total energy can be supplied by a 12 V, 60 $\text{A} \cdot \text{h}$ battery if its internal resistance is negligible? (b) What volume (in liters) of gasoline has a total heat of combustion equal to the energy obtained in part (a)? (See Section 17.6; the density of gasoline is $900\ \text{kg/m}^3$.) (c) If a generator with an average electrical power output of 0.45 kW is connected to the battery, how much time will be required for it to charge the battery fully?

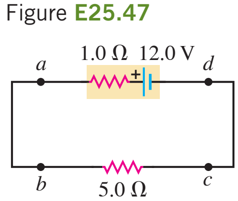
25.44 • A battery has emf $\mathcal{E}$ and internal resistance $r = 2.00\ \Omega$. A 12.0 Ω resistor is connected to the battery, and the resistor consumes electrical power at a rate of 96.0 J/s. What is the emf of the battery?

25.45 •• A 25.0 Ω bulb is connected across the terminals of a 12.0 V battery having 3.50 Ω of internal resistance. What percentage of the power of the battery is dissipated across the internal resistance and hence is not available to the bulb?

25.46 •• A typical small flashlight contains two batteries, each having an emf of 1.5 V, connected in series with a bulb having resistance 17 Ω. (a) If the internal resistance of the batteries is negligible, what power is delivered to the bulb? (b) If the batteries last for 5.0 h, what is the total energy delivered to the bulb? (c) The resistance of real batteries increases as they run down. If the initial internal resistance is negligible, what is the combined internal resistance of both batteries when the power to the bulb has decreased to half its initial value? (Assume that the resistance of the bulb is constant. Actually, it will change somewhat when the current through the filament changes, because this changes the temperature of the filament and hence the resistivity of the filament wire.)

25.47 • In the circuit in **Fig. E25.47**, find (a) the rate of conversion of internal (chemical) energy to electrical energy within the battery; (b) the rate of dissipation of electrical energy in the battery; (c) the rate of dissipation of electrical energy in the external resistor.

Figure **E25.47**



25.48 • A "540 W" electric heater is designed to operate from 120 V lines. (a) What is its operating resistance? (b) What current does it draw? (c) If the line voltage drops to 110 V, what power does the heater take? (Assume that the resistance is constant. Actually, it will change because of the change in temperature.) (d) The heater coils are metallic, so that the resistance of the heater decreases with decreasing temperature. If the change of resistance with temperature is taken into account, will the electrical power consumed by the heater be larger or smaller than what you calculated in part (c)? Explain.

Section 25.6 Theory of Metallic Conduction

25.49 •• Pure silicon at room temperature contains approximately 1.0×10^{16} free electrons per cubic meter. (a) Referring to Table 25.1, calculate the mean free time τ for silicon at room temperature. (b) Your answer in part (a) is much greater than the mean free time for copper given in Example 25.11. Why, then, does pure silicon have such a high resistivity compared to copper?

PROBLEMS

25.50 •• A cell phone or computer battery has three ratings marked on it: a charge capacity listed in mAh (milliamp-hours), an energy capacity in Wh (watt-hours), and a potential rating in volts. (a) What are these three values for your cell phone? (b) Convert the charge capacity Q into coulombs. (c) Convert the energy capacity U into joules. (d) Multiply the charge rating Q by the potential rating V, and verify that this is equivalent to the energy capacity U. (e) If the charge Q were stored on a parallel-plate capacitor with air as the dielectric, at the potential V, what would be the corresponding capacitance? (f) If the energy in the battery were used to heat 1 L of water, estimate the corresponding change in the water temperature? (The heat capacity of water is 4190 $\text{J/kg} \cdot \text{K}$.)

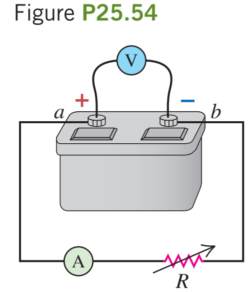
25.51 • An electrical conductor designed to carry large currents has a circular cross section 2.50 mm in diameter and is 14.0 m long. The resistance between its ends is 0.104 Ω. (a) What is the resistivity of the material? (b) If the electric-field magnitude in the conductor is 1.28 V/m, what is the total current? (c) If the material has 8.5×10^{28} free electrons per cubic meter, find the average drift speed under the conditions of part (b).

25.52 •• (a) Estimate the maximum volume of water the hot-water heater in your home can hold. (b) How much heat would be required to raise the temperature of that water from 20°C to a standard household hot-water temperature of 45°C? (Water has a density of 1.00 kg/L and a heat capacity of 4190 $\text{J/kg} \cdot \text{K}$.) (c) Suppose the water should be fully heated in 1.5 h. To what power output does this correspond? (d) If the element has a potential difference of 220 V, what current is required? (e) What should be the resistance of the element?

25.53 •• On your first day at work as an electrical technician, you are asked to determine the resistance per meter of a long piece of wire. The company you work for is poorly equipped. You find a battery, a voltmeter, and an ammeter, but no meter for directly measuring resistance (an ohmmeter). You put the leads from the voltmeter across the terminals of the battery, and the meter reads 12.6 V. You cut off a 20.0 m length of wire and connect it to the battery, with an ammeter in series with it to measure the current in the wire. The ammeter reads 7.00 A. You then cut off a 40.0 m length of wire and connect it to the battery, again with the ammeter in series to measure the current. The ammeter reads 4.20 A. Even though the equipment you have available to you is limited, your boss assures you of its high quality: The ammeter has very small resistance, and the voltmeter has very large resistance. What is the resistance of 1 meter of wire?

25.54 •• In the circuit shown in **Fig. P25.54**, R is a *variable resistor* whose value ranges from 0 to ∞, and a and b are the terminals of a battery that has an emf $\mathcal{E} = 15.0$ V and an internal resistance of 4.00 Ω. The ammeter and voltmeter are idealized meters. As R varies over its full range of values, what will be the largest and smallest readings of (a) the voltmeter and (b) the ammeter? (c) Sketch qualitative graphs of the readings of both meters as functions of R.

Figure **P25.54**



25.55 • A 3.00 m length of copper wire at 20°C has a 1.20-m-long section with diameter 1.60 mm and a 1.80-m-long section with diameter 0.80 mm. There is a current of 2.5 mA in the 1.60-mm-diameter section. (a) What is the current in the 0.80-mm-diameter section? (b) What is the magnitude of $\vec{E}$ in the 1.60-mm-diameter section? (c) What is the magnitude of $\vec{E}$ in the 0.80-mm-diameter section? (d) What is the potential difference between the ends of the 3.00 m length of wire?

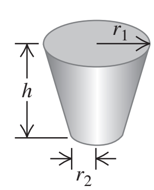
25.56 •• A heating element made of tungsten wire is connected to a large battery that has negligible internal resistance. When the heating element reaches 80.0°C, it consumes electrical energy at a rate of 480 W. What is its power consumption when its temperature is 150.0°C? Assume that the temperature coefficient of resistivity has the value given in Table 25.2 and that it is constant over the temperature range in this problem. In Eq. (25.12) take T_0 to be 20.0°C.

25.57 •• CP BIO **Struck by Lightning.** Lightning strikes can involve currents as high as 25,000 A that last for about 40 μs. If a person is struck by a bolt of lightning with these properties, the current will pass through his body. We shall assume that his mass is 75 kg, that he is wet (after all, he is in a rainstorm) and therefore has a resistance of 1.0 kΩ, and that his body is all water (which is reasonable for a rough, but plausible, approximation). (a) By how many degrees Celsius would this lightning bolt increase the temperature of 75 kg of water? (b) Given that the internal body temperature is about 37°C, would the person's temperature actually increase that much? Why not? What would happen first?

25.58 •• A resistor with resistance R is connected to a battery that has emf 12.0 V and internal resistance $r = 0.40\ \Omega$. For what two values of R will the power dissipated in the resistor be 80.0 W?

25.59 • CALC A material of resistivity ρ is formed into a solid, truncated cone of height h and radii r_1 and r_2 at either end (**Fig. P25.59**). (a) Calculate the resistance of the cone between the two flat end faces. (*Hint:* Imagine slicing the cone into very many thin disks, and calculate the resistance of one such disk.) (b) Show that your result agrees with Eq. (25.10) when $r_1 = r_2$.

Figure **P25.59**



25.60 • CALC The region between two concentric conducting spheres with radii a and b is filled with a conducting material with resistivity ρ. (a) Show that the resistance between the spheres is given by

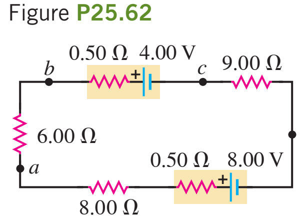
$$R = \frac{\rho}{4\pi}\left(\frac{1}{a} - \frac{1}{b}\right)$$

(b) Derive an expression for the current density as a function of radius, in terms of the potential difference V_{ab} between the spheres. (c) Show that the result in part (a) reduces to Eq. (25.10) when the separation $L = b - a$ between the spheres is small.

25.61 • The potential difference across the terminals of a battery is 8.40 V when there is a current of 1.50 A in the battery from the negative to the positive terminal. When the current is 3.50 A in the reverse direction, the potential difference becomes 10.20 V. (a) What is the internal resistance of the battery? (b) What is the emf of the battery?

25.62 • (a) What is the potential difference V_{ad} in the circuit of **Fig. P25.62**? (b) What is the terminal voltage of the 4.00 V battery? (c) A battery with emf 10.30 V and internal resistance 0.50 Ω is inserted in the circuit at d, with its negative terminal connected to the negative terminal of the 8.00 V battery. What is the difference of potential V_{bc} between the terminals of the 4.00 V battery now?

Figure **P25.62**



25.63 • BIO The average bulk resistivity of the human body (apart from surface resistance of the skin) is about 5.0 $\Omega \cdot$m. The conducting path between the hands can be represented approximately as a cylinder 1.6 m long and 0.10 m in diameter. The skin resistance can be made negligible by soaking the hands in salt water. (a) What is the resistance between the hands if the skin resistance is negligible? (b) What potential difference between the hands is needed for a lethal shock current of 100 mA? (Note that your result shows that small potential differences produce dangerous currents when the skin is damp.) (c) With the current in part (b), what power is dissipated in the body?

25.64 •• BIO A person with body resistance between his hands of 10 kΩ accidentally grasps the terminals of a 14 kV power supply. (a) If the internal resistance of the power supply is 2000 Ω, what is the current through the person's body? (b) What is the power dissipated in his body? (c) If the power supply is to be made safe by increasing its internal resistance, what should the internal resistance be for the maximum current in the above situation to be 1.00 mA or less?

25.65 • A typical cost for electrical power is \$0.120 per kilowatt-hour. (a) Some people leave their porch light on all the time. What is the yearly cost to keep a 75 W bulb burning day and night? (b) Suppose your refrigerator uses 400 W of power when it's running, and it runs 8 hours a day. What is the yearly cost of operating your refrigerator?

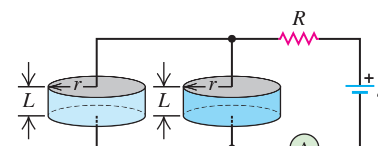
25.66 •• A cylindrical copper cable 1.50 km long is connected across a 220.0 V potential difference. (a) What should be its diameter so that it produces heat at a rate of 90.0 W? (b) What is the electric field inside the cable under these conditions?

25.67 • CALC A 1.50 m cylinder of radius 1.10 cm is made of a complicated mixture of materials. Its resistivity depends on the distance x from the left end and obeys the formula $\rho(x) = a + bx^2$, where a and b are constants. At the left end, the resistivity is $2.25 \times 10^{-8}\ \Omega \cdot$m, while at the right end it is $8.50 \times 10^{-8}\ \Omega \cdot$m. (a) What is the resistance of this rod? (b) What is the electric field at its midpoint if it carries a 1.75 A current? (c) If we cut the rod into two 75.0 cm halves, what is the resistance of each half?

25.68 •• **Compact Fluorescent Bulbs.** Compact fluorescent bulbs are much more efficient at producing light than are ordinary incandescent bulbs. They initially cost much more, but they last far longer and use much less electricity. According to one study of these bulbs, a compact bulb that produces as much light as a 100 W incandescent bulb uses only 23 W of power. The compact bulb lasts 10,000 hours, on the average, and costs \$11.00, whereas the incandescent bulb costs only \$0.75, but lasts just 750 hours. The study assumed that electricity costs \$0.080 per kilowatt-hour and that the bulbs are on for 4.0 h per day. (a) What is the total cost (including the price of the bulbs) to run each bulb for 3.0 years? (b) How much do you save over 3.0 years if you use a compact fluorescent bulb instead of an incandescent bulb? (c) What is the resistance of a "100 W" fluorescent bulb? (Remember, it actually uses only 23 W of power and operates across 120 V.)

25.69 ••• CP Two cylindrical cans with insulating sides and conducting end caps are filled with water, attached to the circuitry shown in **Fig. P25.69**, and used to determine salinity levels. The cans are identical, with radius $r = 5.00$ cm and length $L = 3.00$ cm. The battery supplies a potential of 10.0 V, has a negligible internal resistance, and is connected in series with a resistor $R = 15.0\ \Omega$. The left cylinder is filled with pure distilled water, which has infinite resistivity. The right cylinder is filled with a saltwater solution. It is known that

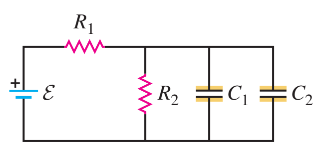
Figure **P25.69**



the resistivity of the saltwater solution is determined by the relationship $\rho = (s_0/s)\ \Omega \cdot \text{m}$, where s is the salinity in parts per thousand (ppt) and $s_0 = 6.30$ ppt. (a) The ammeter registers a current of 484 mA. What is the salinity of the saltwater solution? (b) The left cylinder acts as a capacitor. Use Eq. (24.19) for its capacitance. How much charge is present on its upper plate? Note that pure water has a dielectric constant of 80.4. (c) At what rate is energy dissipated by the saltwater? (d) For what salinity level would the 15.0 Ω resistor dissipate half the power supplied by the battery?

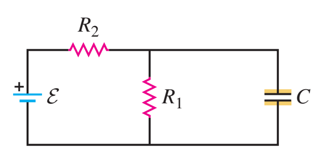
25.70 •• CP Consider the circuit shown in **Fig. P25.70**. The battery has emf 72.0 V and negligible internal resistance. $R_2 = 2.00\ \Omega$, $C_1 = 3.00\ \mu\text{F}$, and $C_2 = 6.00\ \mu\text{F}$. After the capacitors have attained their final charges, the charge on C_1 is $Q_1 = 18.0\ \mu\text{C}$. What is (a) the final charge on C_2; (b) the resistance R_1?

Figure **P25.70**



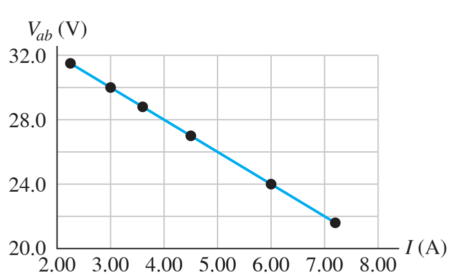
25.71 •• CP Consider the circuit shown in **Fig. P25.71**. The emf source has negligible internal resistance. The resistors have resistances $R_1 = 6.00\ \Omega$ and $R_2 = 4.00\ \Omega$. The capacitor has capacitance $C = 9.00\ \mu\text{F}$. When the capacitor is fully charged, the magnitude of the charge on its plates is $Q = 36.0\ \mu\text{C}$. Calculate the emf $\mathcal{E}$.

Figure **P25.71**



25.72 •• DATA An external resistor R is connected between the terminals of a battery. The value of R varies. For each R value, the current I in the circuit and the terminal voltage V_{ab} of the battery are measured. The results are plotted in **Fig. P25.72**, a graph of V_{ab} versus I that shows the best straight-line fit to the data. (a) Use the graph in Fig. P25.72 to calculate the battery's emf and internal resistance. (b) For what value of R is V_{ab} equal to 80.0% of the battery emf?

Figure **P25.72**



25.73 •• DATA The voltage drop V_{ab} across each of resistors A and B was measured as a function of the current I in the resistor. The results are shown in the table:

Resistor *A*				
I (A)	0.50	1.00	2.00	4.00
V_{ab} (V)	2.55	3.11	3.77	4.58
Resistor *B*				
I (A)	0.50	1.00	2.00	4.00
V_{ab} (V)	1.94	3.88	7.76	15.52

(a) For each resistor, graph V_{ab} as a function of I and graph the resistance $R = V_{ab}/I$ as a function of I. (b) Does resistor A obey Ohm's law? Explain. (c) Does resistor B obey Ohm's law? Explain. (d) What is the power dissipated in A if it is connected to a 4.00 V battery that has negligible internal resistance? (e) What is the power dissipated in B if it is connected to the battery?

25.74 •• DATA According to the U.S. National Electrical Code, copper wire used for interior wiring of houses, hotels, office buildings, and industrial plants is permitted to carry no more than a specified maximum amount of current. The table shows values of the maximum current I_{max} for several common sizes of wire with varnished cambric insulation. The "wire gauge" is a standard used to describe the diameter of wires. Note that the larger the diameter of the wire, the *smaller* the wire gauge.

Wire gauge	Diameter (cm)	I_{max} (A)
14	0.163	18
12	0.205	25
10	0.259	30
8	0.326	40
6	0.412	60
5	0.462	65
4	0.519	85

(a) What considerations determine the maximum current-carrying capacity of household wiring? (b) A total of 4200 W of power is to be supplied through the wires of a house to the household electrical appliances. If the potential difference across the group of appliances is 120 V, determine the gauge of the thinnest permissible wire that can be used. (c) Suppose the wire used in this house is of the gauge found in part (b) and has total length 42.0 m. At what rate is energy dissipated in the wires? (d) The house is built in a community where the consumer cost of electrical energy is \$0.11 per kilowatt-hour. If the house were built with wire of the next larger diameter than that found in part (b), what would be the savings in electricity costs in one year? Assume that the appliances are kept on for an average of 12 hours a day.

25.75 ••• CP CALC A material with resistivity ρ is formed into a cylinder of length L and outer radius r_{outer}. A cylindrical core with radius r_{inner} is removed from the axis of this cylinder and filled with a conducting material, which is attached to a wire. The outer surface of the cylinder is coated with a conducting material and attached to another wire. (a) If the second wire has potential V greater than the first wire, in what direction does the local electric field point inside of the cylinder? (b) The magnitude of this electric field is c/r, where c is a constant and r is the distance from the axis of the cylinder. Use the relationship $V = \int \vec{E} \cdot d\vec{l}$ to determine the constant c. (c) What is the resistance of this device? (d) A 1.00-cm-long hollow cylindrical resistor has an inner radius of 1.50 mm and an outer radius of 3.00 mm. The material is a blend of powdered carbon and ceramic whose resistivity ρ may be altered by changing the amount of carbon. If this device should have a resistance of 6.80 kΩ, what value of ρ should be selected?

25.76 ••• An incandescent light bulb uses a coiled filament of tungsten that is 580 mm long with a diameter of 46.0 μm. At 20.0°C tungsten has a resistivity of $5.25 \times 10^{-8}\ \Omega \cdot \text{m}$. Its temperature coefficient of resistivity is $0.0045\ (\text{C}°)^{-1}$, and this remains accurate even at high temperatures. The temperature of the filament increases linearly with current, from 20°C when no current flows to 2520°C at 1.00 A of current. (a) What is the resistance of the light bulb at 20°C? (b) What is the current through the light bulb when the potential difference across its terminals is 120 V? (*Hint:* First determine the temperature as a function of the current; then use this to determine the resistance as a function of the current. Substitute this result into the equation $V = IR$ and solve for the current I.) (c) What is the resistance when the potential is 120 V? (d) How much energy does the light bulb dissipate in 1 min when 120 V is supplied across its terminals? (e) How much energy does the light bulb dissipate in 1 min when half that voltage is supplied?

CHALLENGE PROBLEMS

25.77 ••• CALC The resistivity of a semiconductor can be modified by adding different amounts of impurities. A rod of semiconducting material of length L and cross-sectional area A lies along the x-axis between $x = 0$ and $x = L$. The material obeys Ohm's law, and its resistivity varies along the rod according to $\rho(x) = \rho_0 \exp(-x/L)$. The end of the rod at $x = 0$ is at a potential V_0 greater than the end at $x = L$. (a) Find the total resistance of the rod and the current in the rod. (b) Find the electric-field magnitude $E(x)$ in the rod as a function of x. (c) Find the electric potential $V(x)$ in the rod as a function of x. (d) Graph the functions $\rho(x)$, $E(x)$, and $V(x)$ for values of x between $x = 0$ and $x = L$.

25.78 ••• An external resistor with resistance R is connected to a battery that has emf $\mathcal{E}$ and internal resistance r. Let P be the electrical power output of the source. By conservation of energy, P is equal to the power consumed by R. What is the value of P in the limit that R is (a) very small; (b) very large? (c) Show that the power output of the battery is a maximum when $R = r$. What is this maximum P in terms of $\mathcal{E}$ and r? (d) A battery has $\mathcal{E} = 64.0$ V and $r = 4.00\ \Omega$. What is the power output of this battery when it is connected to a resistor R, for $R = 2.00\ \Omega$, $R = 4.00\ \Omega$, and $R = 6.00\ \Omega$? Are your results consistent with the general result that you derived in part (b)?

MCAT-STYLE PASSAGE PROBLEMS

BIO **Spiderweb Conductivity**. Some types of spiders build webs that consist of threads made of dry silk coated with a solution of a variety of compounds. This coating leaves the threads, which are used to capture prey, *hygroscopic*—that is, they attract water from the atmosphere. It has been hypothesized that this aqueous coating makes the threads good electrical conductors. To test the electrical properties of coated thread, researchers placed a 5 mm length of thread between two electrical contacts.* The researchers stretched the thread in 1 mm increments to more than twice its original length, and then allowed it to return to its original length, again in 1 mm increments. Some of the resistance measurements are shown in the table:

Resistance of thread ($10^9\ \Omega$)	9	19	41	63	102	76	50	24
Length of thread (mm)	5	7	9	11	13	9	7	5

*Based on F. Vollrath and D. Edmonds, "Consequences of electrical conductivity in an orb spider's capture web," *Naturwissenschaften* (100:12, December 2013, pp. 1163–69).

25.79 What is the best explanation for the behavior exhibited in the data? (a) Longer threads can carry more current than shorter threads do and so make better electrical conductors. (b) The thread stops being a conductor when it is stretched to 13 mm, due to breaks that occur in the thin coating. (c) As the thread is stretched, the coating thins and its resistance increases; as the thread is relaxed, the coating returns nearly to its original state. (d) The resistance of the thread increases with distance from the end of the thread.

25.80 If the conductivity of the thread results from the aqueous coating only, how does the cross-sectional area A of the coating compare when the thread is 13 mm long versus the starting length of 5 mm? Assume that the resistivity of the coating remains constant and the coating is uniform along the thread. $A_{13\text{ mm}}$ is about (a) $\frac{1}{10}A_{5\text{ mm}}$; (b) $\frac{1}{4}A_{5\text{ mm}}$; (c) $\frac{2}{5}A_{5\text{ mm}}$; (d) the same as $A_{5\text{ mm}}$.

25.81 What is the maximum current that flows in the thread during this experiment if the voltage source is a 9 V battery? (a) About 1 A; (b) about 0.1 A; (c) about 1 μA; (d) about 1 nA.

25.82 In another experiment, a piece of the web is suspended so that it can move freely. When either a positively charged object or a negatively charged object is brought near the web, the thread is observed to move toward the charged object. What is the best interpretation of this observation? The web is (a) a negatively charged conductor; (b) a positively charged conductor; (c) either a positively or negatively charged conductor; (d) an electrically neutral conductor.

ANSWERS

Chapter Opening Question ?

(iii) The current out equals the current in. In other words, charge must enter the bulb at the same rate as it exits the bulb. It is not "used up" or consumed as it flows through the bulb.

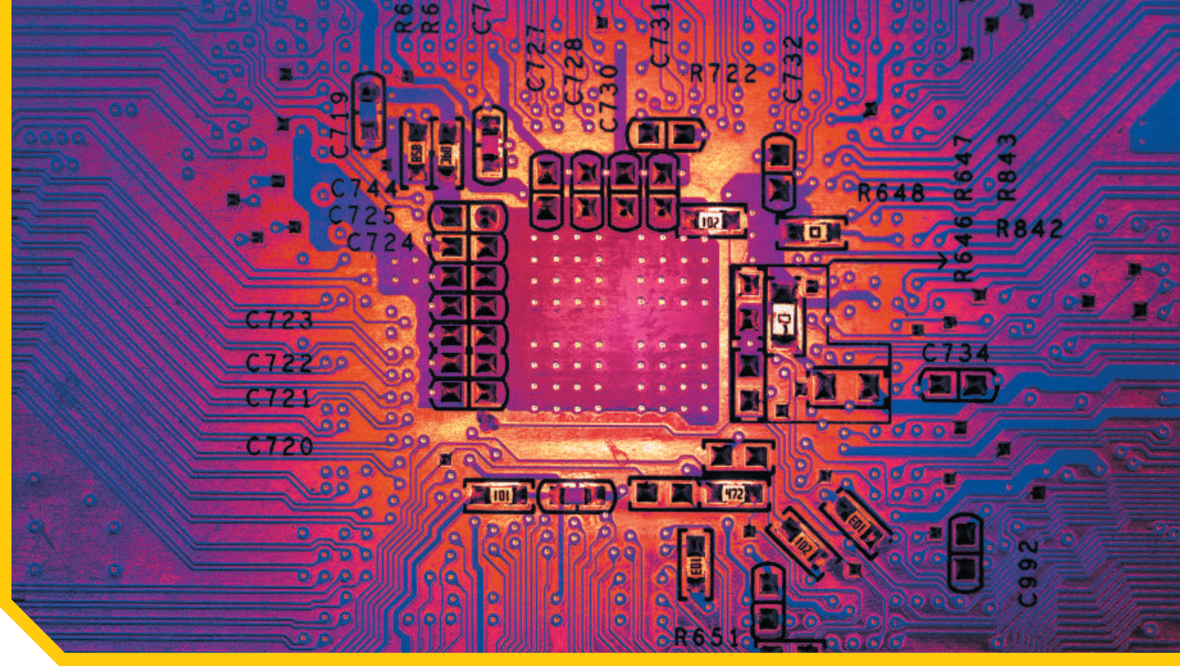
Key Example VARIATION Problems

VP25.3.1 (a) 4.14×10^{-2} V/m (b) 1.21 Ω (c) $7.97 \times 10^{-7}\ \text{m}^2$
VP25.3.2 (a) 0.596 Ω (b) $2.44 \times 10^{-8}\ \Omega \cdot \text{m}$
VP25.3.3 (a) $\frac{1}{2}$ (b) 2 (c) $\frac{1}{2}$ (d) $\frac{1}{2}$
VP25.3.4 (a) 0.625 Ω at 20.0°C, 0.750 Ω at 80.0°C (b) 0.00333 $(\text{C}°)^{-1}$
VP25.5.1 (a) 0.549 A (b) 8.40 V (c) 8.40 V
VP25.5.2 (a) 12.0 V (b) 1.3 Ω
VP25.5.3 (a) 0.0650 A (b) 3.1 Ω
VP25.5.4 (a) 0.590 Ω (b) 1.18 V
VP25.9.1 (a) 1.24 A (b) 29.8 W (c) 2.01 W (d) 27.8 W (e) 27.8 W
VP25.9.2 (a) 0.738 A (b) 6.54 W (c) 8.86 V (d) 0.2 Ω
VP25.9.3 (a) 11.5 V (b) 0.7 Ω (c) 8.29 W (d) 8.64 W (e) 0.3 W
VP25.9.4 (a) $i_2/i_1 = (r + R)/(r + 2R)$ (b) $P_2/P_1 = 2(r + R)^2/(r + 2R)^2$ (c) R_1, R_1

Bridging Problem

(a) 237°C
(b) 162 W initially, 148 W at 1.23 A

? In a complex circuit like the one on this circuit board, is it possible to connect several resistors with different resistances so that all of them have the same potential difference? (i) Yes, and the current will be the same through all of the resistors; (ii) yes, but the current may be different through different resistors; (iii) no; (iv) the answer depends on the value of the potential difference.



26 Direct-Current Circuits

If you look inside a mobile phone, a computer, or under the hood of a car, you'll find circuits of much greater complexity than the simple circuits we studied in Chapter 25. Whether connected by wires or integrated in a semiconductor chip, these circuits often include several sources, resistors, and other circuit elements interconnected in a *network*.

In this chapter we study general methods for analyzing such networks, including how to find voltages and currents of circuit elements. We'll learn how to determine the equivalent resistance for several resistors connected in series or in parallel. For more general networks we need two rules called *Kirchhoff's rules*. One is based on the principle of conservation of charge applied to a junction; the other is derived from energy conservation for a charge moving around a closed loop. We'll discuss instruments for measuring various electrical quantities. We'll also look at a circuit containing resistance and capacitance, in which the current varies with time.

Our principal concern in this chapter is with **direct-current** (dc) circuits, in which the direction of the current does not change with time. Flashlights and automobile wiring systems are examples of direct-current circuits. Household electrical power is supplied in the form of **alternating current** (ac), in which the current oscillates back and forth. The same principles for analyzing networks apply to both kinds of circuits, and we conclude this chapter with a look at household wiring systems. We'll discuss alternating-current circuits in detail in Chapter 31.

LEARNING OUTCOMES

In this chapter, you'll learn...

26.1 How to analyze circuits with multiple resistors in series or parallel.
26.2 Rules that you can apply to any circuit with more than one loop.
26.3 How to use an ammeter, voltmeter, ohmmeter, or potentiometer in a circuit.
26.4 How to analyze circuits that include both a resistor and a capacitor.
26.5 How electric power is distributed in the home.

You'll need to review...

24.2 Capacitors in series and parallel.
25.4 Current, ammeters, and voltmeters.
25.5 Power in a circuit.

26.1 RESISTORS IN SERIES AND PARALLEL

Resistors turn up in all kinds of circuits, ranging from hair dryers and space heaters to circuits that limit or divide current or reduce or divide a voltage. Such circuits often contain several resistors, so it's appropriate to consider *combinations* of resistors. A simple example is a string of light bulbs used for holiday decorations; each bulb acts as a resistor, and from a circuit-analysis perspective the string of bulbs is simply a combination of resistors.

Suppose we have three resistors with resistances R_1, R_2, and R_3. **Figure 26.1** shows four different ways in which they might be connected between points a and b. When several circuit elements such as resistors, batteries, and motors are connected in sequence as in Fig. 26.1a, with only a single current path between the points, we say that they are connected in **series.** We studied *capacitors* in series in Section 24.2; we found that, because of conservation of charge, capacitors in series all have the same charge if they are initially uncharged. In circuits we're often more interested in the *current*, which is charge flow per unit time.

The resistors in Fig. 26.1b are said to be connected in **parallel** between points a and b. Each resistor provides an alternative path between the points. For circuit elements that are connected in parallel, the *potential difference* is the same across each element. We studied capacitors in parallel in Section 24.2.

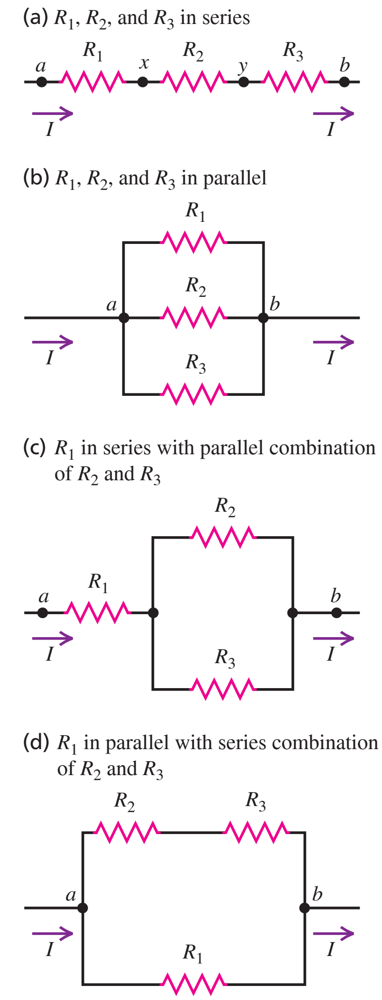
In Fig. 26.1c, resistors R_2 and R_3 are in parallel, and this combination is in series with R_1. In Fig. 26.1d, R_2 and R_3 are in series, and this combination is in parallel with R_1.

For any combination of resistors we can always find a *single* resistor that could replace the combination and result in the same total current and potential difference. For example, a string of holiday light bulbs could be replaced by a single, appropriately chosen light bulb that would draw the same current and have the same potential difference between its terminals as the original string of bulbs. The resistance of this single resistor is called the **equivalent resistance** of the combination. If any one of the networks in Fig. 26.1 were replaced by its equivalent resistance R_{eq}, we could write

$$V_{ab} = IR_{eq} \quad \text{or} \quad R_{eq} = \frac{V_{ab}}{I}$$

where V_{ab} is the potential difference between terminals a and b of the network and I is the current at point a or b. To compute an equivalent resistance, we assume a potential difference V_{ab} across the actual network, compute the corresponding current I, and take the ratio V_{ab}/I.

Figure **26.1** Four different ways of connecting three resistors.



Resistors in Series

We can derive general equations for the equivalent resistance of a series or parallel combination of resistors. In Fig. 26.1a, the three resistors are in *series,* so the current I is the same in all of them. (Recall from Section 25.4 that current is *not* "used up" as it passes through a circuit.) Applying $V = IR$ to each resistor, we have

$$V_{ax} = IR_1 \qquad V_{xy} = IR_2 \qquad V_{yb} = IR_3$$

The potential differences across each resistor need not be the same (except for the special case in which all three resistances are equal). The potential difference V_{ab} across the entire combination is the sum of these individual potential differences:

$$V_{ab} = V_{ax} + V_{xy} + V_{yb} = I(R_1 + R_2 + R_3)$$

and so

$$\frac{V_{ab}}{I} = R_1 + R_2 + R_3$$

The ratio V_{ab}/I is, by definition, the equivalent resistance R_{eq}. Therefore

$$R_{eq} = R_1 + R_2 + R_3$$

It is easy to generalize this to *any* number of resistors:

Resistors in series:

$$R_{eq} = R_1 + R_2 + R_3 + \cdots \tag{26.1}$$

Equivalent resistance of series combination

Resistances of individual resistors

The equivalent resistance of a series combination equals the sum of the individual resistances. The equivalent resistance is *greater than* any individual resistance.

CAUTION **Resistors vs. capacitors in series** Don't confuse *resistors* in series with *capacitors* in series. Resistors in series add *directly* [Eq. (26.1)] because the voltage across each is directly proportional to its resistance and to the common current. Capacitors in series [Eq. (24.5)] add *reciprocally;* the voltage across each is directly proportional to the common charge but *inversely* proportional to the individual capacitance.

Resistors in Parallel

If the resistors are in *parallel,* as in Fig. 26.1b, the current through each resistor need not be the same. But the potential difference between the terminals of each resistor must be the same and equal to V_{ab} (**Fig. 26.2**). (Remember that the potential difference between any two points does not depend on the path taken between the points.) Let's call the currents in the three resistors I_1, I_2, and I_3. Then from $I = V/R$,

$$I_1 = \frac{V_{ab}}{R_1} \qquad I_2 = \frac{V_{ab}}{R_2} \qquad I_3 = \frac{V_{ab}}{R_3}$$

Figure **26.2** A car's headlights and taillights are connected in parallel. Hence each light is exposed to the full potential difference supplied by the car's electrical system, giving maximum brightness. Another advantage is that if one headlight or taillight burns out, the other one keeps shining (see Example 26.2).

In general, the current is different through each resistor. Because charge neither accumulates at nor drains out of point a, the total current I must equal the sum of the three currents in the resistors:

$$I = I_1 + I_2 + I_3 = V_{ab}\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right)$$

or

$$\frac{I}{V_{ab}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

But by the definition of the equivalent resistance R_{eq}, $I/V_{ab} = 1/R_{eq}$, so

$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

Again it is easy to generalize to *any* number of resistors in parallel:

Resistors in parallel:

$$\underbrace{\frac{1}{R_{eq}}}_{\text{Equivalent resistance of parallel combination}} = \underbrace{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}}_{\text{Resistances of individual resistors}} + \cdots \qquad (26.2)$$

The reciprocal of the equivalent resistance of a parallel combination equals the sum of the reciprocals of the individual resistances. The equivalent resistance is always *less than* any individual resistance.

CAUTION **Resistors vs. capacitors in parallel** Note the differences between *resistors* in parallel and *capacitors* in parallel. Resistors in parallel add *reciprocally* [Eq. (26.2)] because the current in each is proportional to the common voltage across them and *inversely* proportional to the resistance of each. Capacitors in parallel add *directly* [Eq. (24.7)] because the charge on each is proportional to the common voltage across them and *directly* proportional to the capacitance of each.

For the special case of *two* resistors in parallel,

$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{R_1 + R_2}{R_1 R_2} \quad \text{and}$$

$$R_{eq} = \frac{R_1 R_2}{R_1 + R_2} \qquad \text{(two resistors in parallel)} \qquad (26.3)$$

Because $V_{ab} = I_1R_1 = I_2R_2$, it follows that

$$\frac{I_1}{I_2} = \frac{R_2}{R_1} \quad \text{(two resistors in parallel)} \tag{26.4}$$

Thus the currents carried by two resistors in parallel are *inversely proportional* to their resistances. More current goes through the path of least resistance.

PROBLEM-SOLVING STRATEGY 26.1 Resistors in Series and Parallel

IDENTIFY *the relevant concepts:* As in Fig. 26.1, many resistor networks are made up of resistors in series, in parallel, or a combination thereof. Such networks can be replaced by a single equivalent resistor. The logic is similar to that of Problem-Solving Strategy 24.1 for networks of capacitors.

SET UP *the problem* using the following steps:

1. Make a drawing of the resistor network.
2. Identify groups of resistors connected in series or parallel.
3. Identify the target variables. They could include the equivalent resistance of the network, the potential difference across each resistor, or the current through each resistor.

EXECUTE *the solution* as follows:

1. Use Eq. (26.1) or (26.2), respectively, to find the equivalent resistance for series or parallel combinations.
2. If the network is more complex, try reducing it to series and parallel combinations. For example, in Fig. 26.1c we first replace the parallel combination of R_2 and R_3 with its equivalent resistance; this then forms a series combination with R_1. In Fig. 26.1d, the combination of R_2 and R_3 in series forms a parallel combination with R_1.
3. Keep in mind that the total potential difference across resistors connected in series is the sum of the individual potential differences. The potential difference across resistors connected in parallel is the same for every resistor and equals the potential difference across the combination.
4. The current through resistors connected in series is the same through every resistor and equals the current through the combination. The total current through resistors connected in parallel is the sum of the currents through the individual resistors.

EVALUATE *your answer:* Check whether your results are consistent. The equivalent resistance of resistors connected in series should be greater than that of any individual resistor; that of resistors in parallel should be less than that of any individual resistor.

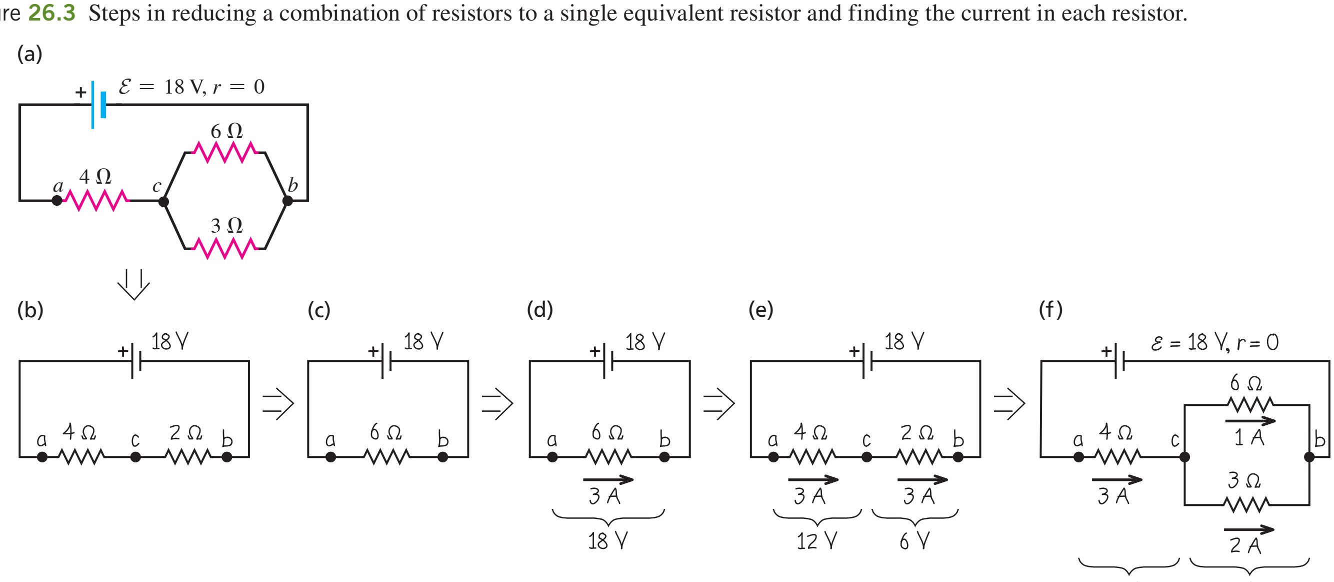
EXAMPLE 26.1 Equivalent resistance

WITH VARIATION PROBLEMS

Find the equivalent resistance of the network in **Fig. 26.3a** and the current in each resistor. The source of emf has negligible internal resistance.

IDENTIFY and SET UP This network of three resistors is a *combination* of series and parallel resistances, as in Fig. 26.1c. We determine the equivalent resistance of the parallel 6 Ω and 3 Ω resistors, and then that of their series combination with the 4 Ω resistor: This is the equivalent resistance R_{eq} of the network as a whole. We then find the current in the emf, which is the same as that in the 4 Ω resistor. The potential difference is the same across each of the parallel 6 Ω and 3 Ω resistors; we use this to determine how the current is divided between these.

Figure 26.3 Steps in reducing a combination of resistors to a single equivalent resistor and finding the current in each resistor.



Continued

EXECUTE Figures 26.3b and 26.3c show successive steps in reducing the network to a single equivalent resistance R_{eq}. From Eq. (26.2), the $6\ \Omega$ and $3\ \Omega$ resistors in parallel in Fig. 26.3a are equivalent to the single $2\ \Omega$ resistor in Fig. 26.3b:

$$\frac{1}{R_{6\,\Omega+3\,\Omega}} = \frac{1}{6\ \Omega} + \frac{1}{3\ \Omega} = \frac{1}{2\ \Omega}$$

[Equation (26.3) gives the same result.] From Eq. (26.1) the series combination of this $2\ \Omega$ resistor with the $4\ \Omega$ resistor is equivalent to the single $6\ \Omega$ resistor in Fig. 26.3c.

We reverse these steps to find the current in each resistor of the original network. In the circuit shown in Fig. 26.3d (identical to Fig. 26.3c), the current is $I = V_{ab}/R = (18\ \text{V})/(6\ \Omega) = 3\ \text{A}$. So the current in the $4\ \Omega$ and $2\ \Omega$ resistors in Fig. 26.3e (identical to Fig. 26.3b) is also 3 A. The potential difference V_{cb} across the $2\ \Omega$ resistor is therefore $V_{cb} = IR = (3\ \text{A})(2\ \Omega) = 6\ \text{V}$. This potential difference must also be 6 V in Fig. 26.3f (identical to Fig. 26.3a). From $I = V_{cb}/R$, the currents in the $6\ \Omega$ and $3\ \Omega$ resistors in Fig. 26.3f are, respectively, $(6\ \text{V})/(6\ \Omega) = 1\ \text{A}$ and $(6\ \text{V})/(3\ \Omega) = 2\ \text{A}$.

EVALUATE Note that for the two resistors in parallel between points c and b in Fig. 26.3f, there is twice as much current through the $3\ \Omega$ resistor as through the $6\ \Omega$ resistor; more current goes through the path of least resistance, in accordance with Eq. (26.4). Note also that the total current through these two resistors is 3 A, the same as it is through the $4\ \Omega$ resistor between points a and c.

KEYCONCEPT When resistors are combined in *series,* the current is the same through each resistor but the potential differences across the resistors may not be the same. The equivalent resistance of a series combination equals the sum of the individual resistances. When resistors are combined in *parallel,* the potential difference across each resistor is the same but the currents through the resistors may not be the same. The reciprocal of the equivalent resistance of a parallel combination equals the sum of the reciprocals of the individual resistances.

EXAMPLE 26.2 Series versus parallel combinations

WITH VARIATION PROBLEMS

Two identical incandescent light bulbs, each with resistance $R = 2\ \Omega$, are connected to a source with $\mathcal{E} = 8\ \text{V}$ and negligible internal resistance. Find the current through each bulb, the potential difference across each bulb, and the power delivered to each bulb and to the entire network if the bulbs are connected (a) in series and (b) in parallel. (c) Suppose one of the bulbs burns out; that is, its filament breaks and current can no longer flow through it. What happens to the other bulb in the series case? In the parallel case?

IDENTIFY and SET UP The light bulbs are just resistors in simple series and parallel connections (**Figs. 26.4a** and 26.4b). Once we find the current I through each bulb, we can find the power delivered to each bulb by using Eq. (25.18), $P = I^2R = V^2/R$.

EXECUTE (a) From Eq. (26.1) the equivalent resistance of the two bulbs between points a and c in Fig. 26.4a is $R_{eq} = 2R = 2(2\ \Omega) = 4\ \Omega$. In series, the current is the same through each bulb:

$$I = \frac{V_{ac}}{R_{eq}} = \frac{8\ \text{V}}{4\ \Omega} = 2\ \text{A}$$

Since the bulbs have the same resistance, the potential difference is the same across each bulb:

$$V_{ab} = V_{bc} = IR = (2\ \text{A})(2\ \Omega) = 4\ \text{V}$$

From Eq. (25.18), the power delivered to each bulb is

$$P = I^2R = (2\ \text{A})^2(2\ \Omega) = 8\ \text{W} \qquad \text{or}$$

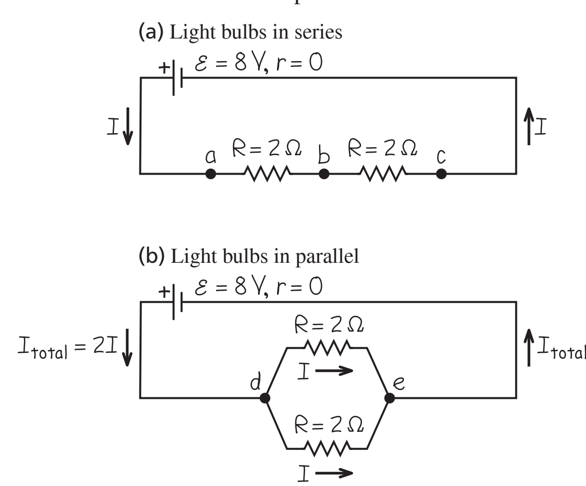
$$P = \frac{V_{ab}^2}{R} = \frac{V_{bc}^2}{R} = \frac{(4\ \text{V})^2}{2\ \Omega} = 8\ \text{W}$$

The total power delivered to both bulbs is $P_{tot} = 2P = 16\ \text{W}$.

(b) If the bulbs are in parallel, as in Fig. 26.4b, the potential difference V_{de} across each bulb is the same and equal to 8 V, the terminal voltage of the source. Hence the current through each light bulb is

$$I = \frac{V_{de}}{R} = \frac{8\ \text{V}}{2\ \Omega} = 4\ \text{A}$$

Figure 26.4 Our sketches for this problem.



and the power delivered to each bulb is

$$P = I^2R = (4\ \text{A})^2(2\ \Omega) = 32\ \text{W} \qquad \text{or}$$

$$P = \frac{V_{de}^2}{R} = \frac{(8\ \text{V})^2}{2\ \Omega} = 32\ \text{W}$$

Both the potential difference across each bulb and the current through each bulb are twice as great as in the series case. Hence the power delivered to each bulb is *four* times greater, and each bulb is brighter.

The total power delivered to the parallel network is $P_{total} = 2P = 64\ \text{W}$, four times greater than in the series case. The increased power compared to the series case isn't obtained "for free"; energy is extracted from the source four times more rapidly in the parallel case than in the series case. If the source is a battery, it will be used up four times as fast.

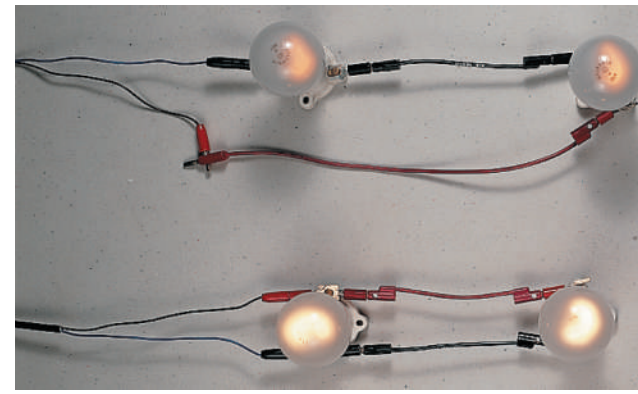
(c) In the series case the same current flows through both bulbs. If one bulb burns out, there will be no current in the circuit, and neither bulb will glow.

In the parallel case the potential difference across either bulb is unchanged if a bulb burns out. The current through the functional bulb and the power delivered to it are unchanged.

EVALUATE Our calculation isn't completely accurate, because the resistance $R = V/I$ of real light bulbs depends on the potential difference V across the bulb. That's because the filament resistance increases with increasing operating temperature and therefore with increasing V. But bulbs connected in series across a source do in fact glow less brightly than when connected in parallel across the same source (**Fig. 26.5**).

KEYCONCEPT Two or more resistors connected in parallel to a source each have more current through them and each draw more power than if they were connected in series to the same source.

Figure 26.5 When connected to the same source, two incandescent light bulbs in series (shown at top) draw less power and glow less brightly than when they are in parallel (shown at bottom).



TEST YOUR UNDERSTANDING OF SECTION 26.1 Suppose all three of the resistors shown in Fig. 26.1 have the same resistance, so $R_1 = R_2 = R_3 = R$. Rank the four arrangements shown in parts (a)–(d) of Fig. 26.1 in order of their equivalent resistance, from highest to lowest.

ANSWER

❙ **(a), (c), (d), (b)** Here's why: The three resistors in Fig. 26.1a are in series, so $R_{eq} = R + R + R = 3R$. In Fig. 26.1b the three resistors are in parallel, so $1/R_{eq} = 1/R + 1/R + 1/R = 3/R$ and $R_{eq} = R/3$. In Fig. 26.1c the second and third resistors are in parallel, so their equivalent resistance R_{23} is given by $1/R_{23} = 1/R + 1/R = 2/R$; hence $R_{23} = R/2$. This combination is in series with the first resistor, so the three resistors together have equivalent resistance $R_{eq} = R + R/2 = 3R/2$. In Fig. 26.1d the second and third resistors are in series, so their equivalent resistance is $R_{23} = R + R = 2R$. This combination is in parallel with the first resistor, so the equivalent resistance of the three-resistor combination is given by $1/R_{eq} = 1/R + 1/2R = 3/2R$. Hence $R_{eq} = 2R/3$.

26.2 KIRCHHOFF'S RULES

Many practical resistor networks cannot be reduced to simple series-parallel combinations. **Figure 26.6a** shows a dc power supply with emf $\mathcal{E}_1$ charging a battery with a smaller emf $\mathcal{E}_2$ and feeding current to a light bulb with resistance R. Figure 26.6b is a "bridge" circuit, used in many different types of measurement and control systems. (Problem 26.74 describes one important application of a "bridge" circuit.) To analyze these networks, we'll use the techniques developed by the German physicist Gustav Robert Kirchhoff (1824–1887).

First, here are two terms that we'll use often. A **junction** in a circuit is a point where three or more conductors meet. A **loop** is any closed conducting path. In Fig. 26.6a points a and b are junctions, but points c and d are not; in Fig. 26.6b points a, b, c, and d are junctions, but points e and f are not. The blue lines in Figs. 26.6a and 26.6b show some possible loops in these circuits.

Kirchhoff's rules are the following two statements:

Kirchhoff's junction rule (valid at any junction): The sum of the currents into any junction ...

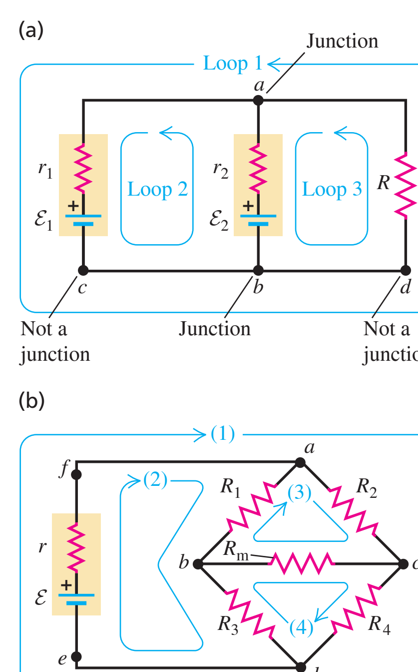
$$\sum I = 0 \quad \text{... equals zero.} \tag{26.5}$$

Kirchhoff's loop rule (valid for any closed loop): The sum of the potential differences around any loop ...

$$\sum V = 0 \quad \text{... equals zero.} \tag{26.6}$$

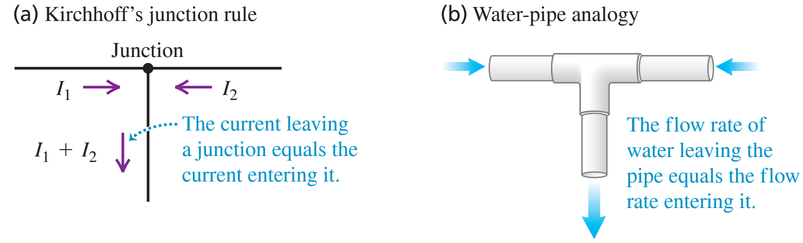
Note that the potential differences V in Eq. (26.6) include those associated with all circuit elements in the loop, including emfs and resistors.

Figure 26.6 Two networks that cannot be reduced to simple series-parallel combinations of resistors.



CAUTION Current need not split equally at a junction As a rule, at a junction where the current in a wire splits to follow two or more separate paths, the current does *not* split equally among the different paths. More current flows along a path that has less resistance.

Figure 26.7 Kirchhoff's junction rule states that as much current flows into a junction as flows out of it.



The junction rule is based on *conservation of electric charge*. No charge can accumulate at a junction, so the total charge entering the junction per unit time must equal the total charge leaving per unit time (**Fig. 26.7a**). Charge per unit time is current, so if we consider the currents entering a junction to be positive and those leaving to be negative, the algebraic sum of currents into a junction must be zero. It's like a T branch in a water pipe (Fig. 26.7b); if you have a total of 1 liter per minute coming in the two pipes, you can't have 3 liters per minute going out the third pipe. We used the junction rule (without saying so) in Section 26.1 in the derivation of Eq. (26.2) for resistors in parallel.

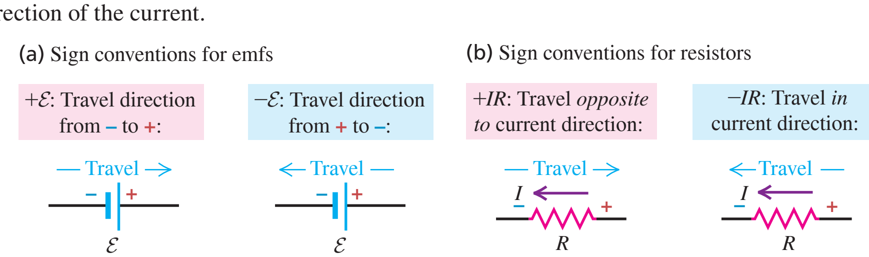
The loop rule is a statement that the electrostatic force is *conservative*. Suppose we go around a loop, measuring potential differences across successive circuit elements as we go. When we return to the starting point, we must find that the *algebraic sum* of these differences is zero; otherwise, we could not say that the potential at this point has a definite value.

Sign Conventions for the Loop Rule

In applying the loop rule, we need some sign conventions. Problem-Solving Strategy 26.2 describes in detail how to use these, but here's a quick overview. We first assume a direction for the current in each branch of the circuit and mark it on a diagram of the circuit. Then, starting at any point in the circuit, we imagine traveling around a loop, adding emfs and IR terms as we come to them. When we travel through a source in the direction from $-$ to $+$, the emf is considered to be *positive;* when we travel from $+$ to $-$, the emf is considered to be *negative* (**Fig. 26.8a**). When we travel through a resistor in the *same* direction as the assumed current, the IR term is *negative* because the current goes in the direction of decreasing potential. When we travel through a resistor in the direction *opposite* to the assumed current, the IR term is *positive* because this represents a rise of potential (Fig. 26.8b).

Kirchhoff's two rules are all we need to solve a wide variety of network problems. Usually, some of the emfs, currents, and resistances are known, and others are unknown. We must always obtain from Kirchhoff's rules a number of independent equations equal to the number of unknowns so that we can solve the equations simultaneously. Often the hardest part of the solution is keeping track of algebraic signs!

Figure 26.8 Use these sign conventions when you apply Kirchhoff's loop rule. In each part of the figure "Travel" is the direction that we imagine going around the loop, which is not necessarily the direction of the current.



PROBLEM-SOLVING STRATEGY 26.2 Kirchhoff's Rules

IDENTIFY *the relevant concepts:* Kirchhoff's rules are useful for analyzing any electric circuit.

SET UP *the problem* using the following steps:

1. Draw a circuit diagram, leaving room to label all quantities, known and unknown. Indicate an assumed direction for each unknown current and emf. (Kirchhoff's rules will yield the magnitudes *and directions* of unknown currents and emfs. If the actual direction of a quantity is opposite to your assumption, the resulting quantity will have a negative sign.)
2. As you label currents, it's helpful to use Kirchhoff's junction rule, as in **Fig. 26.9**, so as to express the currents in terms of as few quantities as possible.
3. Identify the target variables.

EXECUTE *the solution* as follows:

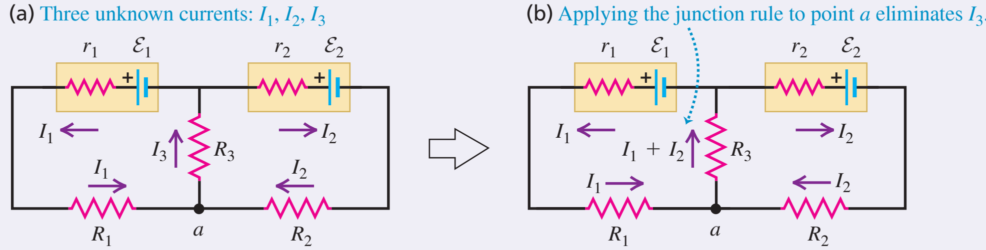
1. Choose any loop in the network and choose a direction (clockwise or counterclockwise) to travel around the loop as you apply Kirchhoff's loop rule. The direction need not be the same as any assumed current direction.
2. Travel around the loop in the chosen direction, adding potential differences algebraically as you cross them. Use the sign conventions of Fig. 26.8.
3. Equate the sum obtained in step 2 to zero in accordance with the loop rule.
4. If you need more independent equations, choose another loop and repeat steps 1–3; continue until you have as many independent equations as unknowns or until every circuit element has been included in at least one loop.
5. Solve the equations simultaneously to determine the unknowns.
6. You can use the loop-rule bookkeeping system to find the potential V_{ab} of any point a with respect to any other point b. Start at b and add the potential changes you encounter in going from b to a; use the same sign rules as in step 2. The algebraic sum of these changes is $V_{ab} = V_a - V_b$.

EVALUATE *your answer:* Check all the steps in your algebra. Apply steps 1 and 2 to a loop you have not yet considered; if the sum of potential drops isn't zero, you've made an error somewhere.

Figure **26.9** Applying the junction rule to point a reduces the number of unknown currents from three to two.

(a) Three unknown currents: I_1, I_2, I_3

(b) Applying the junction rule to point a eliminates I_3.



EXAMPLE 26.3 A single-loop circuit

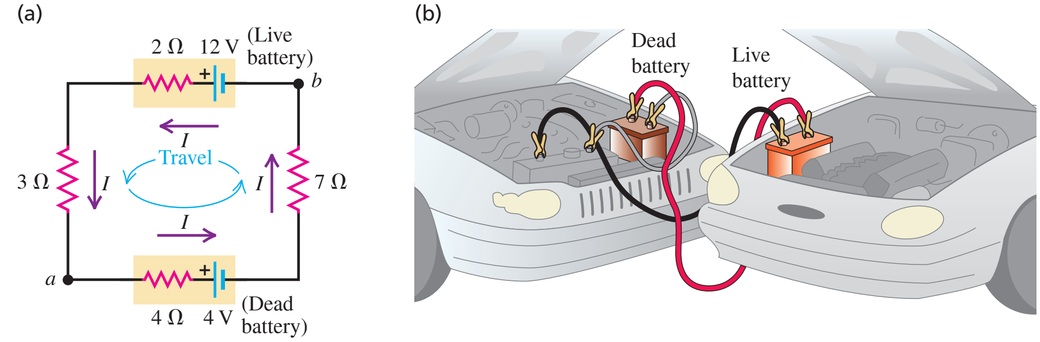
WITH VARIATION PROBLEMS

The circuit shown in **Fig. 26.10a** contains two batteries, each with an emf and an internal resistance, and two resistors. Find (a) the current in the circuit, (b) the potential difference V_{ab}, and (c) the power output of the emf of each battery.

IDENTIFY and SET UP There are no junctions in this single-loop circuit, so we don't need Kirchhoff's junction rule. To apply Kirchhoff's loop rule, we first assume a direction for the current; let's assume a counterclockwise direction as shown in Fig. 26.10a.

Figure **26.10** (a) In this example we travel around the loop in the same direction as the assumed current, so all the IR terms are negative. The potential decreases as we travel from + to − through the bottom emf but increases as we travel from − to + through the top emf. (b) A real-life example of a circuit of this kind.

(a) (b)



Continued

EXECUTE (a) Starting at a and traveling counterclockwise with the current, we add potential increases and decreases and equate the sum to zero as in Eq. (26.6):

$$-I(4\ \Omega) - 4\ \text{V} - I(7\ \Omega) + 12\ \text{V} - I(2\ \Omega) - I(3\ \Omega) = 0$$

Collecting like terms and solving for I, we find

$$8\ \text{V} = I(16\ \Omega) \qquad \text{and} \qquad I = 0.5\ \text{A}$$

The positive result for I shows that our assumed current direction is correct.

(b) To find V_{ab}, the potential at a with respect to b, we start at b and add potential changes as we go toward a. There are two paths from b to a; taking the lower one, we find

$$V_{ab} = (0.5\ \text{A})(7\ \Omega) + 4\ \text{V} + (0.5\ \text{A})(4\ \Omega) = 9.5\ \text{V}$$

Point a is at 9.5 V higher potential than b. All the terms in this sum, including the IR terms, are positive because each represents an *increase* in potential as we go from b to a. For the upper path,

$$V_{ab} = 12\ \text{V} - (0.5\ \text{A})(2\ \Omega) - (0.5\ \text{A})(3\ \Omega) = 9.5\ \text{V}$$

Here the IR terms are negative because our path goes in the direction of the current, with potential decreases through the resistors. The results for V_{ab} are the same for both paths, as they must be in order for the total potential change around the loop to be zero.

(c) The power outputs of the emf of the 12 V and 4 V batteries are

$$P_{12\text{V}} = \mathcal{E}I = (12\ \text{V})(0.5\ \text{A}) = 6\ \text{W}$$
$$P_{4\text{V}} = \mathcal{E}I = (-4\ \text{V})(0.5\ \text{A}) = -2\ \text{W}$$

The negative sign in $\mathcal{E}$ for the 4 V battery appears because the current actually runs from the higher-potential side of the battery to the lower-potential side. The negative value of P means that we are *storing* energy in that battery; the 12 V battery is *recharging* it (if it is in fact rechargeable; otherwise, we're destroying it).

EVALUATE By applying the expression $P = I^2R$ to each of the four resistors in Fig. 26.10a, you can show that the total power dissipated in all four resistors is 4 W. Of the 6 W provided by the emf of the 12 V battery, 2 W goes into storing energy in the 4 V battery and 4 W is dissipated in the resistances.

The circuit shown in Fig. 26.10a is much like that used when a fully charged 12 V storage battery (in a car with its engine running) "jump-starts" a car with a run-down battery (Fig. 26.10b). The run-down battery is slightly recharged in the process. The 3 Ω and 7 Ω resistors in Fig. 26.10a represent the resistances of the jumper cables and of the conducting path through the car with the run-down battery. (The values of the resistances in actual automobiles and jumper cables are considerably lower, and the emf of a run-down car battery isn't much less than 12 V.)

KEYCONCEPT In any loop of a circuit, Kirchhoff's loop rule applies: The sum of the potential changes as you travel around the loop must be zero. If the circuit has only a single loop, the current has the same value everywhere in the circuit.

EXAMPLE 26.4 Charging a battery

WITH VARIATION PROBLEMS

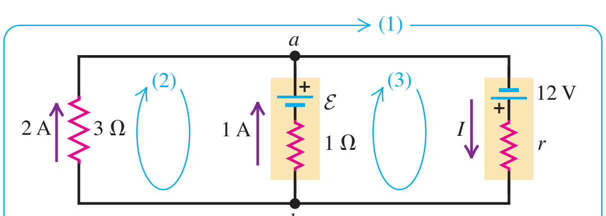
In the circuit shown in **Fig. 26.11**, a 12 V power supply with unknown internal resistance r is connected to a run-down rechargeable battery with unknown emf $\mathcal{E}$ and internal resistance 1 Ω and to an indicator light bulb of resistance 3 Ω carrying a current of 2 A. The current through the run-down battery is 1 A in the direction shown. Find r, $\mathcal{E}$, and the current I through the power supply.

IDENTIFY and SET UP This circuit has more than one loop, so we must apply both the junction and loop rules. We assume the direction of the current through the 12 V power supply, and the polarity of the run-down battery, to be as shown in Fig. 26.11. There are three target variables, so we need three equations.

EXECUTE We apply the junction rule, Eq. (26.5), to point a:

$$-I + 1\ \text{A} + 2\ \text{A} = 0 \qquad \text{so} \qquad I = 3\ \text{A}$$

Figure 26.11 In this circuit a power supply charges a run-down battery and lights a bulb. An assumption has been made about the polarity of the emf $\mathcal{E}$ of the battery. Is this assumption correct?



To determine r, we apply the loop rule, Eq. (26.6), to the large, outer loop (1):

$$12\ \text{V} - (3\ \text{A})r - (2\ \text{A})(3\ \Omega) = 0 \qquad \text{so} \qquad r = 2\ \Omega$$

To determine $\mathcal{E}$, we apply the loop rule to the left-hand loop (2):

$$-\mathcal{E} + (1\ \text{A})(1\ \Omega) - (2\ \text{A})(3\ \Omega) = 0 \qquad \text{so} \qquad \mathcal{E} = -5\ \text{V}$$

The negative value for $\mathcal{E}$ shows that the actual polarity of this emf is opposite to that shown in Fig. 26.11. As in Example 26.3, the battery is being recharged.

EVALUATE Try applying the junction rule at point b instead of point a, and try applying the loop rule counterclockwise rather than clockwise around loop (1). You'll get the same results for I and r. We can check our result for $\mathcal{E}$ by using loop (3):

$$12\ \text{V} - (3\ \text{A})(2\ \Omega) - (1\ \text{A})(1\ \Omega) + \mathcal{E} = 0$$

which again gives us $\mathcal{E} = -5$ V.

As an additional check, we note that $V_{ba} = V_b - V_a$ equals the voltage across the 3 Ω resistance, which is $(2\ \text{A})(3\ \Omega) = 6$ V. Going from a to b by the right-hand branch, we encounter potential differences $+12\ \text{V} - (3\ \text{A})(2\ \Omega) = +6$ V, and going by the middle branch, we find $-(-5\ \text{V}) + (1\ \text{A})(1\ \Omega) = +6$ V. The three ways of getting V_{ba} give the same results.

KEYCONCEPT In any circuit that has more than one loop, Kirchhoff's junction rule applies: At each junction, the sum of the currents into the junction must be zero.

EXAMPLE 26.5 Power in a battery-charging circuit

WITH VARIATION PROBLEMS

In the circuit of Example 26.4 (shown in Fig. 26.11), find the power delivered by the 12 V power supply and by the battery being recharged, and find the power dissipated in each resistor.

IDENTIFY and SET UP We use the results of Section 25.5, in which we found that the power delivered *from* an emf to a circuit is $\mathcal{E}I$ and the power delivered *to* a resistor from a circuit is $V_{ab}I = I^2R$. We know the values of all relevant quantities from Example 26.4.

EXECUTE The power output from the emf of the power supply is

$$P_{\text{supply}} = \mathcal{E}_{\text{supply}}I_{\text{supply}} = (12\text{ V})(3\text{ A}) = 36\text{ W}$$

The power dissipated in the power supply's internal resistance r is

$$P_{r-\text{supply}} = I_{\text{supply}}{}^2 r_{\text{supply}} = (3\text{ A})^2(2\ \Omega) = 18\text{ W}$$

so the power supply's *net* power output is $P_{\text{net}} = 36\text{ W} - 18\text{ W} = 18\text{ W}$. Alternatively, from Example 26.4 the terminal voltage of the battery is $V_{ba} = 6\text{ V}$, so the net power output is

$$P_{\text{net}} = V_{ba}I_{\text{supply}} = (6\text{ V})(3\text{ A}) = 18\text{ W}$$

The power output of the emf $\mathcal{E}$ of the battery being charged is

$$P_{\text{emf}} = \mathcal{E}I_{\text{battery}} = (-5\text{ V})(1\text{ A}) = -5\text{ W}$$

This is negative because the 1 A current runs through the battery from the higher-potential side to the lower-potential side. (As we mentioned in Example 26.4, the polarity assumed for this battery in Fig. 26.11 was wrong.) We are storing energy in the battery as we charge it. Additional power is dissipated in the battery's internal resistance; this power is

$$P_{r-\text{battery}} = I_{\text{battery}}{}^2 r_{\text{battery}} = (1\text{ A})^2(1\ \Omega) = 1\text{ W}$$

The total power input to the battery is thus $1\text{ W} + |-5\text{ W}| = 6\text{ W}$. Of this, 5 W represents useful energy stored in the battery; the remainder is wasted in its internal resistance. The power dissipated in the light bulb is

$$P_{\text{bulb}} = I_{\text{bulb}}{}^2 R_{\text{bulb}} = (2\text{ A})^2(3\ \Omega) = 12\text{ W}$$

EVALUATE As a check, note that all of the power from the supply is accounted for. Of the 18 W of net power from the power supply, 5 W goes to recharge the battery, 1 W is dissipated in the battery's internal resistance, and 12 W is dissipated in the light bulb.

KEYCONCEPT In any circuit that contains a source (such as a battery or power supply) and resistors, the sum of the power dissipated in each resistor must equal the power supplied by the source.

EXAMPLE 26.6 A complex network

WITH VARIATION PROBLEMS

Figure 26.12 shows a "bridge" circuit of the type described at the beginning of this section (see Fig. 26.6b). Find the current in each resistor and the equivalent resistance of the network of five resistors.

IDENTIFY and SET UP This network is neither a series combination nor a parallel combination. Hence we must use Kirchhoff's rules to find the values of the target variables. There are five unknown currents, but by applying the junction rule to junctions a and b, we can represent them in terms of three unknown currents I_1, I_2, and I_3, as shown in Fig. 26.12.

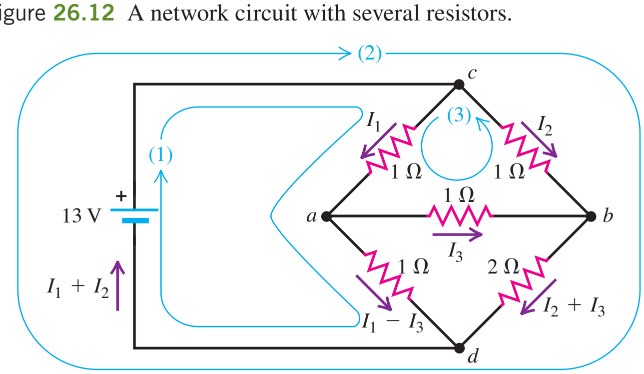
EXECUTE We apply the loop rule to the three loops shown:

$$13\text{ V} - I_1(1\ \Omega) - (I_1 - I_3)(1\ \Omega) = 0 \qquad (1)$$

$$-I_2(1\ \Omega) - (I_2 + I_3)(2\ \Omega) + 13\text{ V} = 0 \qquad (2)$$

$$-I_1(1\ \Omega) - I_3(1\ \Omega) + I_2(1\ \Omega) = 0 \qquad (3)$$

Figure **26.12** A network circuit with several resistors.



One way to solve these simultaneous equations is to solve Eq. (3) for I_2, obtaining $I_2 = I_1 + I_3$, and then substitute this expression into Eq. (2) to eliminate I_2. We then have

$$13\text{ V} = I_1(2\ \Omega) - I_3(1\ \Omega) \qquad (1')$$

$$13\text{ V} = I_1(3\ \Omega) + I_3(5\ \Omega) \qquad (2')$$

Now we can eliminate I_3 by multiplying Eq. (1′) by 5 and adding the two equations. We obtain

$$78\text{ V} = I_1(13\ \Omega) \qquad I_1 = 6\text{ A}$$

We substitute this result into Eq. (1′) to obtain $I_3 = -1\text{ A}$, and from Eq. (3) we find $I_2 = 5\text{ A}$. The negative value of I_3 tells us that its direction is opposite to the direction we assumed.

The total current through the network is $I_1 + I_2 = 11\text{ A}$, and the potential drop across it is equal to the battery emf, 13 V. The equivalent resistance of the network is therefore

$$R_{\text{eq}} = \frac{13\text{ V}}{11\text{ A}} = 1.2\ \Omega$$

EVALUATE You can check our results for I_1, I_2, and I_3 by substituting them back into Eqs. (1)–(3). What do you find?

KEYCONCEPT Some circuits have combinations of resistors that are neither in series nor in parallel. To analyze any such circuit, use Kirchhoff's loop rule and junction rule.

EXAMPLE 26.7 A potential difference in a complex network

WITH VARIATION PROBLEMS

In the circuit of Example 26.6 (Fig. 26.12), find the potential difference V_{ab}.

IDENTIFY and SET UP Our target variable $V_{ab} = V_a - V_b$ is the potential at point a with respect to point b. To find it, we start at point b and follow a path to point a, adding potential rises and drops as we go. We can follow any of several paths from b to a; the result must be the same for all such paths, which gives us a way to check our result.

EXECUTE The simplest path is through the center $1\ \Omega$ resistor. In Example 26.6 we found $I_3 = -1$ A, showing that the actual current direction through this resistor is from right to left. Thus, as we go from b to a, there is a *drop* of potential with magnitude $|I_3|R = (1\text{ A})(1\ \Omega) = 1$ V. Hence $V_{ab} = -1$ V, and the potential at a is 1 V less than at point b.

EVALUATE To check our result, let's try a path from b to a that goes through the lower two resistors. The currents through these are

$$I_2 + I_3 = 5\text{ A} + (-1\text{ A}) = 4\text{ A} \quad \text{and}$$
$$I_1 - I_3 = 6\text{ A} - (-1\text{ A}) = 7\text{ A}$$

and so

$$V_{ab} = -(4\text{ A})(2\ \Omega) + (7\text{ A})(1\ \Omega) = -1\text{ V}$$

You can confirm this result by using some other paths from b to a.

KEYCONCEPT A useful way to restate Kirchhoff's loop rule is that for any path in a circuit from point a to point b, the potential difference V_{ab} must be the same.

TEST YOUR UNDERSTANDING OF SECTION 26.2 Subtract Eq. (1) from Eq. (2) in Example 26.6. To which loop in Fig. 26.12 does this equation correspond? Would this equation have simplified the solution of Example 26.6?

ANSWER

loop *cbdac*, no Equation (2) minus Eq. (1) gives $-I_2(1\ \Omega) - (I_2 + I_3)(2\ \Omega) + (I_1 - I_3)(1\ \Omega) + I_1(1\ \Omega) = 0$. We can obtain this equation by applying the loop rule around the path from c to b to d to a to c in Fig. 26.12. This isn't an independent equation, so it would not have helped with the solution of Example 26.6.

Figure 26.13 Both this voltmeter (left) and ammeter (right) are d'Arsonval galvanometers. The difference has to do with their internal connections (see Fig. 26.15).

26.3 ELECTRICAL MEASURING INSTRUMENTS

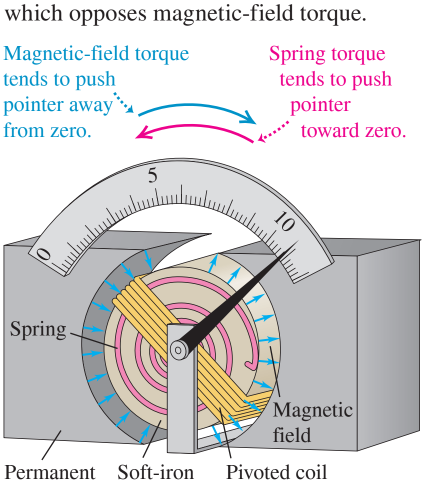
We've been talking about potential difference, current, and resistance for two chapters, so it's about time we said something about how to *measure* these quantities. Many common devices, including car instrument panels, battery chargers, and inexpensive electrical instruments, measure potential difference (voltage), current, or resistance with a **d'Arsonval galvanometer** (**Fig. 26.13**). In the following discussion we'll often call it just a *meter*. A pivoted coil of fine wire is placed in the magnetic field of a permanent magnet (**Fig. 26.14**). Attached to the coil is a spring, similar to the hairspring on the balance wheel of a watch. In the equilibrium position, with no current in the coil, the pointer is at zero. When there is a current in the coil, the magnetic field exerts a torque on the coil that is proportional to the current. (We'll discuss this magnetic interaction in detail in Chapter 27.) As the coil turns, the spring exerts a restoring torque that is proportional to the angular displacement.

Thus the angular deflection of the coil and pointer is directly proportional to the coil current, and the device can be calibrated to measure current. The maximum deflection, typically 90° or so, is called *full-scale deflection*. The essential electrical characteristics of the meter are the current I_{fs} required for full-scale deflection (typically on the order of 10 μA to 10 mA) and the resistance R_c of the coil (typically on the order of 10 Ω to 1000 Ω).

The meter deflection is proportional to the *current* in the coil. If the coil obeys Ohm's law, the current is proportional to the *potential difference* between the terminals of the coil, and the deflection is also proportional to this potential difference. For example, consider a meter whose coil has a resistance $R_c = 20.0\ \Omega$ and that deflects full scale when the current in its coil is $I_{fs} = 1.00$ mA. The corresponding potential difference for full-scale deflection is

$$V = I_{fs}R_c = (1.00 \times 10^{-3}\text{ A})(20.0\ \Omega) = 0.0200\text{ V}$$

Figure 26.14 A d'Arsonval galvanometer, showing a pivoted coil with attached pointer, a permanent magnet supplying a magnetic field that is uniform in magnitude, and a spring to provide restoring torque, which opposes magnetic-field torque.



Ammeters

A current-measuring instrument is usually called an **ammeter** (or milliammeter, microammeter, and so forth, depending on the range). *An ammeter always measures the current passing through it.* An *ideal* ammeter, discussed in Section 25.4, would have *zero* resistance, so including it in a branch of a circuit would not affect the current in that branch.

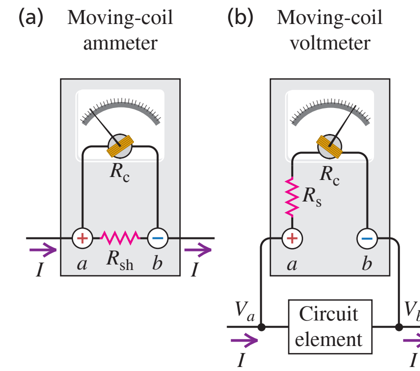
Real ammeters always have a finite resistance, but it is always desirable for an ammeter to have as little resistance as possible.

We can adapt any meter to measure currents that are larger than its full-scale reading by connecting a resistor in parallel with it (**Fig. 26.15a**) so that some of the current bypasses the meter coil. The parallel resistor is called a **shunt resistor** or simply a *shunt,* denoted as R_{sh}.

Suppose we want to make a meter with full-scale current I_{fs} and coil resistance R_c into an ammeter with full-scale reading I_a. To determine the shunt resistance R_{sh} needed, note that at full-scale deflection the total current through the parallel combination is I_a, the current through the coil of the meter is I_{fs}, and the current through the shunt is the difference $I_a - I_{fs}$. The potential difference V_{ab} is the same for both paths, so

$$I_{fs}R_c = (I_a - I_{fs})R_{sh} \qquad \text{(for an ammeter)} \qquad (26.7)$$

Figure **26.15** Using the same meter to measure (a) current and (b) voltage.



EXAMPLE 26.8 Designing an ammeter

What shunt resistance is required to make the 1.00 mA, 20.0 Ω meter described above into an ammeter with a range of 0 to 50.0 mA?

IDENTIFY and SET UP Since the meter is being used as an ammeter, its internal connections are as shown in Fig. 26.15a. Our target variable is the shunt resistance R_{sh}, which we'll find from Eq. (26.7). The ammeter must handle a maximum current $I_a = 50.0 \times 10^{-3}$ A. The coil resistance is $R_c = 20.0\ \Omega$, and the meter shows full-scale deflection when the current through the coil is $I_{fs} = 1.00 \times 10^{-3}$ A.

EXECUTE Solving Eq. (26.7) for R_{sh}, we find

$$R_{sh} = \frac{I_{fs}R_c}{I_a - I_{fs}} = \frac{(1.00 \times 10^{-3}\ \text{A})(20.0\ \Omega)}{(50.0 \times 10^{-3}\ \text{A}) - (1.00 \times 10^{-3}\ \text{A})}$$
$$= 0.408\ \Omega$$

EVALUATE It's useful to consider the equivalent resistance R_{eq} of the ammeter as a whole. From Eq. (26.2),

$$R_{eq} = \left(\frac{1}{R_c} + \frac{1}{R_{sh}}\right)^{-1} = \left(\frac{1}{20.0\ \Omega} + \frac{1}{0.408\ \Omega}\right)^{-1}$$
$$= 0.400\ \Omega$$

The shunt resistance is so small in comparison to the coil resistance that the equivalent resistance is very nearly equal to the shunt resistance. The result is an ammeter with a low equivalent resistance and the desired 0–50.0 mA range. At full-scale deflection, $I = I_a = 50.0$ mA, the current through the galvanometer is 1.00 mA, the current through the shunt resistor is 49.0 mA, and $V_{ab} = 0.0200$ V. If the current I is *less* than 50.0 mA, the coil current and the deflection are proportionally less.

KEYCONCEPT An ammeter measures the current that passes through it. It should have a very small equivalent resistance in order to minimize its effect on the circuit being measured. This can be accomplished using a small shunt resistance in parallel with the meter.

Voltmeters

This same basic meter may also be used to measure potential difference or *voltage.* A voltage-measuring device is called a **voltmeter.** A voltmeter always measures the potential difference between two points, and its terminals must be connected to these points. (Example 25.6 in Section 25.4 described what can happen if a voltmeter is connected incorrectly.) As we discussed in Section 25.4, an ideal voltmeter would have *infinite* resistance, so connecting it between two points in a circuit would not alter any of the currents. Real voltmeters always have finite resistance, but a voltmeter should have large enough resistance that connecting it in a circuit does not change the other currents appreciably.

For the meter of Example 26.8, the voltage across the meter coil at full-scale deflection is only $I_{fs}R_c = (1.00 \times 10^{-3}\ \text{A})(20.0\ \Omega) = 0.0200$ V. We can extend this range by connecting a resistor R_s in *series* with the coil (Fig. 26.15b). Then only a fraction of the total potential difference appears across the coil itself, and the remainder appears across R_s. For a voltmeter with full-scale reading V_V, we need a series resistor R_s in Fig. 26.15b such that

$$V_V = I_{fs}(R_c + R_s) \qquad \text{(for a voltmeter)} \qquad (26.8)$$

BIO APPLICATION
Electromyography A fine needle containing two electrodes is being inserted into a muscle in this patient's hand. By using a sensitive voltmeter to measure the potential difference between these electrodes, a physician can probe the muscle's electrical activity. This is an important technique for diagnosing neurological and neuromuscular diseases.

EXAMPLE 26.9 Designing a voltmeter

What series resistance is required to make the 1.00 mA, 20.0 Ω meter described above into a voltmeter with a range of 0 to 10.0 V?

IDENTIFY and SET UP Since this meter is being used as a voltmeter, its internal connections are as shown in Fig. 26.15b. The maximum allowable voltage across the voltmeter is $V_V = 10.0\text{ V}$. We want this to occur when the current through the coil is $I_{fs} = 1.00 \times 10^{-3}\text{ A}$. Our target variable is the series resistance R_s, which we find from Eq. (26.8).

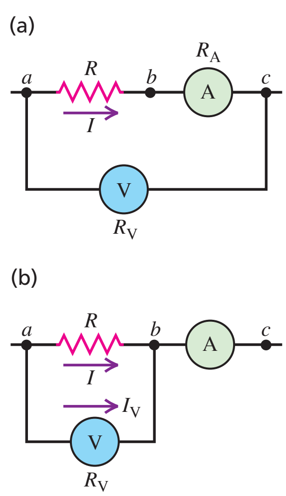
EXECUTE From Eq. (26.8),

$$R_s = \frac{V_V}{I_{fs}} - R_c = \frac{10.0\text{ V}}{0.00100\text{ A}} - 20.0\ \Omega = 9980\ \Omega$$

EVALUATE At full-scale deflection, $V_{ab} = 10.0\text{ V}$, the voltage across the meter is 0.0200 V, the voltage across R_s is 9.98 V, and the current through the voltmeter is 0.00100 A. Most of the voltage appears across the series resistor. The meter's equivalent resistance is a desirably high $R_{eq} = 20.0\ \Omega + 9980\ \Omega = 10{,}000\ \Omega$. Such a meter is called a "1000 ohms-per-volt" meter, referring to the ratio of resistance to full-scale deflection. In normal operation the current through the circuit element being measured (I in Fig. 26.15b) is much greater than 0.00100 A, and the resistance between points a and b in the circuit is much less than 10,000 Ω. The voltmeter draws off only a small fraction of the current and thus disturbs the circuit being measured only slightly.

KEYCONCEPT A voltmeter measures the potential difference (voltage) between its terminals. It should have a very large equivalent resistance in order to minimize the amount of current it draws from the circuit being measured. This can be accomplished using a large shunt resistance in series with the meter.

Figure 26.16 Ammeter–voltmeter method for measuring resistance.



Ammeters and Voltmeters in Combination

A voltmeter and an ammeter can be used together to measure *resistance* and *power*. The resistance R of a resistor equals the potential difference V_{ab} between its terminals divided by the current I; that is, $R = V_{ab}/I$. The power input P to any circuit element is the product of the potential difference across it and the current through it: $P = V_{ab}I$. In principle, the most straightforward way to measure R or P is to measure V_{ab} and I simultaneously.

With practical ammeters and voltmeters this isn't quite as simple as it seems. In **Fig. 26.16a**, ammeter A reads the current I in the resistor R. Voltmeter V, however, reads the *sum* of the potential difference V_{ab} across the resistor and the potential difference V_{bc} across the ammeter. If we transfer the voltmeter terminal from c to b, as in Fig. 26.16b, then the voltmeter reads the potential difference V_{ab} correctly, but the ammeter now reads the *sum* of the current I in the resistor and the current I_V in the voltmeter. Either way, we have to correct the reading of one instrument or the other unless the corrections are small enough to be negligible.

EXAMPLE 26.10 Measuring resistance I

The voltmeter in the circuit of Fig. 26.16a reads 12.0 V and the ammeter reads 0.100 A. The meter resistances are $R_V = 10{,}000\ \Omega$ (for the voltmeter) and $R_A = 2.00\ \Omega$ (for the ammeter). What are the resistance R and the power dissipated in the resistor?

IDENTIFY and SET UP The ammeter reads the current $I = 0.100\text{ A}$ through the resistor, and the voltmeter reads the potential difference between a and c. If the ammeter were *ideal* (that is, if $R_A = 0$), there would be zero potential difference between b and c, the voltmeter reading $V = 12.0\text{ V}$ would be equal to the potential difference V_{ab} across the resistor, and the resistance would be equal to $R = V/I = (12.0\text{ V})/(0.100\text{ A}) = 120\ \Omega$. The ammeter is *not* ideal, however (its resistance is $R_A = 2.00\ \Omega$), so the voltmeter reading V is actually the sum of the potential differences V_{bc} (across the ammeter) and V_{ab} (across the resistor). We use Ohm's law to find the voltage V_{bc} from the known current and ammeter resistance. Then we solve for V_{ab} and R. Given these, we are able to calculate the power P into the resistor.

EXECUTE From Ohm's law, $V_{bc} = IR_A = (0.100\text{ A})(2.00\ \Omega) = 0.200\text{ V}$ and $V_{ab} = IR$. The sum of these is $V = 12.0\text{ V}$, so the potential difference across the resistor is $V_{ab} = V - V_{bc} = (12.0\text{ V}) - (0.200\text{ V}) = 11.8\text{ V}$. Hence the resistance is

$$R = \frac{V_{ab}}{I} = \frac{11.8\text{ V}}{0.100\text{ A}} = 118\ \Omega$$

The power dissipated in this resistor is

$$P = V_{ab}I = (11.8\text{ V})(0.100\text{ A}) = 1.18\text{ W}$$

EVALUATE You can confirm this result for the power by using the alternative formula $P = I^2R$. Do you get the same answer?

KEYCONCEPT To find the resistance of a resistor and the power that the resistor dissipates requires measuring both the current through the resistor and the potential difference across the resistor. To do this, put an ammeter in series with the resistance and put a voltmeter across the resistor and ammeter.

EXAMPLE 26.11 Measuring resistance II

Suppose the meters of Example 26.10 are connected to a different resistor as shown in Fig. 26.16b, and the readings obtained on the meters are the same as in Example 26.10. What is the value of this new resistance R, and what is the power dissipated in the resistor?

IDENTIFY and SET UP In Example 26.10 the ammeter read the actual current through the resistor, but the voltmeter reading was not the same as the potential difference across the resistor. Now the situation is reversed: The voltmeter reading $V = 12.0$ V shows the actual potential difference V_{ab} across the resistor, but the ammeter reading $I_A = 0.100$ A is *not* equal to the current I through the resistor. Applying the junction rule at b in Fig. 26.16b shows that $I_A = I + I_V$, where I_V is the current through the voltmeter. We find I_V from the given values of V and the voltmeter resistance R_V, and we use this value to find the resistor current I. We then determine the resistance R from I and the voltmeter reading, and calculate the power as in Example 26.10.

EXECUTE We have $I_V = V/R_V = (12.0\text{ V})/(10{,}000\ \Omega) = 1.20$ mA. The actual current I in the resistor is $I = I_A - I_V = 0.100\text{ A} - 0.0012\text{ A} = 0.0988$ A, and the resistance is

$$R = \frac{V_{ab}}{I} = \frac{12.0\text{ V}}{0.0988\text{ A}} = 121\ \Omega$$

The power dissipated in the resistor is

$$P = V_{ab}I = (12.0\text{ V})(0.0988\text{ A}) = 1.19\text{ W}$$

EVALUATE Had the meters been ideal, our results would have been $R = 12.0\text{ V}/0.100\text{ A} = 120\ \Omega$ and $P = VI = (12.0\text{ V})(0.100\text{ A}) = 1.2$ W both here and in Example 26.10. The actual (correct) results are not too different in either case. That's because the ammeter and voltmeter are nearly ideal: Compared with the resistance R under test, the ammeter resistance R_A is very small and the voltmeter resistance R_V is very large. Under these conditions, treating the meters as ideal yields pretty good results; accurate work requires calculations as in these two examples.

KEYCONCEPT An alternative way to find the resistance of a resistor and the power that the resistor dissipates is to put a voltmeter across the resistor and put an ammeter in the circuit downstream of the resistor–voltmeter combination.

Ohmmeters

An alternative method for measuring resistance is to use a d'Arsonval meter in an arrangement called an **ohmmeter.** It consists of a meter, a resistor, and a source (often a flashlight battery) connected in series (**Fig. 26.17**). The resistance R to be measured is connected between terminals x and y.

The series resistance R_s is variable; it is adjusted so that when terminals x and y are short-circuited (that is, when $R = 0$), the meter deflects full scale. When nothing is connected to terminals x and y, so that the circuit between x and y is *open* (that is, when $R \rightarrow \infty$), there is no current and hence no deflection. For any intermediate value of R the meter deflection depends on the value of R, and the meter scale can be calibrated to read the resistance R directly. Larger currents correspond to smaller resistances, so this scale reads backward compared to the scale showing the current.

In situations in which high precision is required, instruments containing d'Arsonval meters have been supplanted by electronic instruments with direct digital readouts. Digital voltmeters can be made with extremely high internal resistance, of the order of 100 MΩ. **Figure 26.18** shows a digital *multimeter,* an instrument that can measure voltage, current, or resistance over a wide range.

Figure **26.17** Ohmmeter circuit. The resistor R_s has a variable resistance, as is indicated by the arrow through the resistor symbol. To use the ohmmeter, first connect x directly to y and adjust R_s until the meter reads zero. Then connect x and y across the resistor R and read the scale.

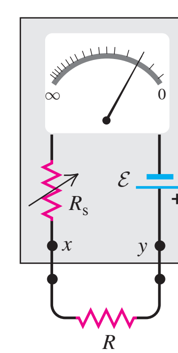


Figure **26.18** This digital multimeter can be used as a voltmeter (red arc), ammeter (yellow arc), or ohmmeter (green arc).

The Potentiometer

The *potentiometer* is an instrument that can be used to measure the emf of a source without drawing any current from the source; it also has a number of other useful applications. Essentially, it balances an unknown potential difference against an adjustable, measurable potential difference.

Figure 26.19 A potentiometer.

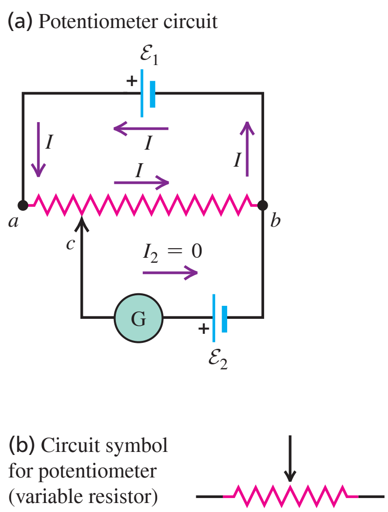


Figure 26.19a shows the principle of the potentiometer. A resistance wire *ab* of total resistance R_{ab} is permanently connected to the terminals of a source of known emf $\mathcal{E}_1$. A sliding contact *c* is connected through the galvanometer G to a second source whose emf $\mathcal{E}_2$ is to be measured. As contact *c* is moved along the resistance wire, the resistance R_{cb} between points *c* and *b* varies; if the resistance wire is uniform, R_{cb} is proportional to the length of wire between *c* and *b*. To determine the value of $\mathcal{E}_2$, contact *c* is moved until a position is found at which the galvanometer shows no deflection; this corresponds to zero current passing through $\mathcal{E}_2$. With $I_2 = 0$, Kirchhoff's loop rule gives

$$\mathcal{E}_2 = IR_{cb}$$

With $I_2 = 0$, the current I produced by the emf $\mathcal{E}_1$ has the same value no matter what the value of the emf $\mathcal{E}_2$. We calibrate the device by replacing $\mathcal{E}_2$ by a source of known emf; then to find any unknown emf $\mathcal{E}_2$, we measure the length of wire *cb* for which $I_2 = 0$. Note: For this to work, V_{ab} must be greater than $\mathcal{E}_2$.

The term *potentiometer* is also used for any variable resistor, usually having a circular resistance element and a sliding contact controlled by a rotating shaft and knob. Figure. 26.19b shows the circuit symbol for a potentiometer.

TEST YOUR UNDERSTANDING OF SECTION 26.3 You want to measure the current through and the potential difference across the 2 Ω resistor shown in Fig. 26.12 (Example 26.6 in Section 26.2). (a) How should you connect an ammeter and a voltmeter to do this? (i) Both ammeter and voltmeter in series with the 2 Ω resistor; (ii) ammeter in series with the 2 Ω resistor and voltmeter connected between points *b* and *d*; (iii) ammeter connected between points *b* and *d* and voltmeter in series with the 2 Ω resistor; (iv) both ammeter and voltmeter connected between points *b* and *d*. (b) What resistances should these meters have? (i) Both ammeter and voltmeter resistances should be much greater than 2 Ω; (ii) ammeter resistance should be much greater than 2 Ω and voltmeter resistance should be much less than 2 Ω; (iii) ammeter resistance should be much less than 2 Ω and voltmeter resistance should be much greater than 2 Ω; (iv) both ammeter and voltmeter resistances should be much less than 2 Ω.

ANSWER

(a) **(ii)**, (b) **(iii)** An ammeter must always be placed in series with the circuit element of interest, and a voltmeter must always be placed in parallel. Ideally the ammeter would have zero resistance and the voltmeter would have infinite resistance so that their presence would have no effect on either the resistor current or the voltage. Neither of these idealizations is possible, but the ammeter resistance should be much less than 2 Ω and the voltmeter resistance should be much greater than 2 Ω.

26.4 *R-C* CIRCUITS

In the circuits we have analyzed up to this point, we have assumed that all the emfs and resistances are *constant* (time independent) so that all the potentials, currents, and powers are also independent of time. But in the simple act of charging or discharging a capacitor we find a situation in which the currents, voltages, and powers *do* change with time.

Many devices incorporate circuits in which a capacitor is alternately charged and discharged. These include flashing traffic lights, automobile turn signals, and electronic flash units. Understanding what happens in such circuits is thus of great practical importance.

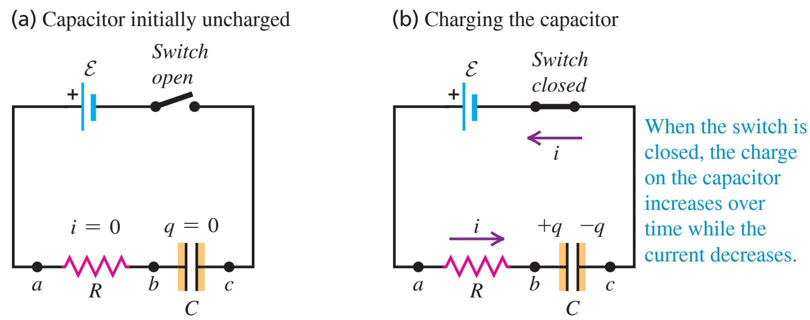


Figure 26.20 Charging a capacitor. (a) Just before the switch is closed, the charge q is zero. (b) When the switch closes (at $t = 0$), the current jumps from zero to $\mathcal{E}/R$. As time passes, q approaches Q_f and the current i approaches zero.

Charging a Capacitor

Figure 26.20 shows a simple circuit for charging a capacitor. A circuit such as this that has a resistor and a capacitor in series is called an ***R-C* circuit.** We idealize the battery (or power supply) to have a constant emf $\mathcal{E}$ and zero internal resistance $(r = 0)$, and we ignore the resistance of all the connecting conductors.

We begin with the capacitor initially uncharged (Fig. 26.20a); then at some initial time $t = 0$ we close the switch, completing the circuit and permitting current around the loop to begin charging the capacitor (Fig. 26.20b). For all practical purposes, the current begins at the same instant in every conducting part of the circuit, and at each instant the current is the same in every part.

Because the capacitor in Fig. 26.20 is initially uncharged, the potential difference v_{bc} across it is zero at $t = 0$. At this time, from Kirchhoff's loop law, the voltage v_{ab} across the resistor R is equal to the battery emf $\mathcal{E}$. The initial $(t = 0)$ current through the resistor, which we'll call I_0, is given by Ohm's law: $I_0 = v_{ab}/R = \mathcal{E}/R$.

As the capacitor charges, its voltage v_{bc} increases and the potential difference v_{ab} across the resistor decreases, corresponding to a decrease in current. The sum of these two voltages is constant and equal to $\mathcal{E}$. After a long time the capacitor is fully charged, the current decreases to zero, and v_{ab} across the resistor becomes zero. Then the entire battery emf $\mathcal{E}$ appears across the capacitor and $v_{bc} = \mathcal{E}$.

Let q represent the charge on the capacitor and i the current in the circuit at some time t after the switch has been closed. We choose the positive direction for the current to correspond to positive charge flowing onto the left-hand capacitor plate, as in Fig. 26.20b. The instantaneous potential differences v_{ab} and v_{bc} are

$$v_{ab} = iR \qquad v_{bc} = \frac{q}{C}$$

Using these in Kirchhoff's loop rule, we find

$$\mathcal{E} - iR - \frac{q}{C} = 0 \tag{26.9}$$

The potential drops by an amount iR as we travel from a to b and by q/C as we travel from b to c. Solving Eq. (26.9) for i, we find

$$i = \frac{\mathcal{E}}{R} - \frac{q}{RC} \tag{26.10}$$

At time $t = 0$, when the switch is first closed, the capacitor is uncharged, and so $q = 0$. Substituting $q = 0$ into Eq. (26.10), we find that the *initial* current I_0 is given by $I_0 = \mathcal{E}/R$, as we have already noted. If the capacitor were not in the circuit, the last term in Eq. (26.10) would not be present; then the current would be *constant* and equal to $\mathcal{E}/R$.

As the charge q increases, the term q/RC becomes larger and the capacitor charge approaches its final value, which we'll call Q_f. The current decreases and eventually becomes zero. When $i = 0$, Eq. (26.10) gives

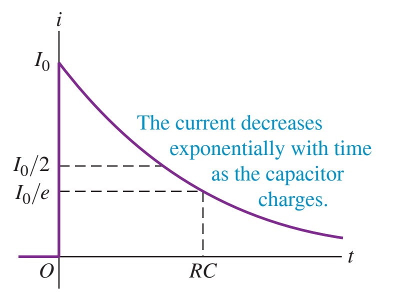
$$\frac{\mathcal{E}}{R} = \frac{Q_f}{RC} \qquad Q_f = C\mathcal{E} \tag{26.11}$$

Note that the final charge Q_f does not depend on R.

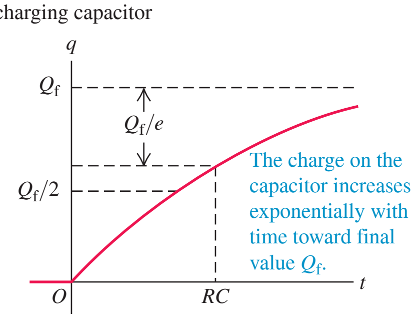
CAUTION Lowercase means time-varying Up to this point we have been working with constant potential differences (voltages), currents, and charges, and we have used *capital* letters V, I, and Q, respectively, to denote these quantities. To distinguish between quantities that vary with time and those that are constant, we'll use *lowercase* letters v, i, and q for time-varying voltages, currents, and charges, respectively. We suggest that you follow this same convention in your own work.

Figure **26.21** Current i and capacitor charge q as functions of time for the circuit of Fig. 26.20. The initial current is I_0 and the initial capacitor charge is zero. The current asymptotically approaches zero, and the capacitor charge asymptotically approaches a final value of Q_f.

(a) Graph of current versus time for a charging capacitor



(b) Graph of capacitor charge versus time for a charging capacitor



BIO APPLICATION Pacemakers and Capacitors This x-ray image shows a pacemaker implanted in a patient with a malfunctioning sinoatrial node, the part of the heart that generates the electrical signal to trigger heartbeats. The pacemaker circuit contains a battery, a capacitor, and a computer-controlled switch. To maintain regular beating, once per second the switch discharges the capacitor and sends an electrical pulse along the lead to the heart. The switch then flips to allow the capacitor to recharge for the next pulse.

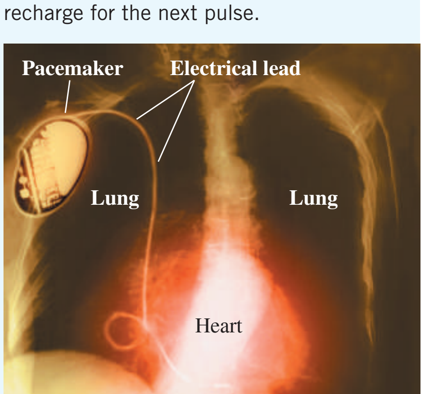


Figure 26.21 shows the current and capacitor charge as functions of time. At the instant the switch is closed $(t = 0)$, the current jumps from zero to its initial value $I_0 = \mathcal{E}/R$; after that, it gradually approaches zero. The capacitor charge starts at zero and gradually approaches the final value given by Eq. (26.11), $Q_f = C\mathcal{E}$.

We can derive general expressions for charge q and current i as functions of time. With our choice of the positive direction for current (Fig. 26.20b), i equals the rate at which positive charge arrives at the left-hand (positive) plate of the capacitor, so $i = dq/dt$. Making this substitution in Eq. (26.10), we have

$$\frac{dq}{dt} = \frac{\mathcal{E}}{R} - \frac{q}{RC} = -\frac{1}{RC}(q - C\mathcal{E})$$

We can rearrange this to

$$\frac{dq}{q - C\mathcal{E}} = -\frac{dt}{RC}$$

and then integrate both sides. We change the integration variables to q' and t' so that we can use q and t for the upper limits. The lower limits are $q' = 0$ and $t' = 0$:

$$\int_0^q \frac{dq'}{q' - C\mathcal{E}} = -\int_0^t \frac{dt'}{RC}$$

When we carry out the integration, we get

$$\ln\left(\frac{q - C\mathcal{E}}{-C\mathcal{E}}\right) = -\frac{t}{RC}$$

Exponentiating both sides (that is, taking the inverse logarithm) and solving for q, we find

$$\frac{q - C\mathcal{E}}{-C\mathcal{E}} = e^{-t/RC}$$

***R-C* circuit, charging capacitor:**

Capacitor charge · Capacitance · Final capacitor charge $= C\mathcal{E}$

$$q = C\mathcal{E}(1 - e^{-t/RC}) = Q_f(1 - e^{-t/RC}) \qquad (26.12)$$

Battery emf · Time since switch closed · Resistance

The instantaneous current i is just the time derivative of Eq. (26.12):

***R-C* circuit, charging capacitor:**

Current · Battery emf · Time since switch closed · Initial current $= \mathcal{E}/R$

$$i = \frac{dq}{dt} = \frac{\mathcal{E}}{R}e^{-t/RC} = I_0 e^{-t/RC} \qquad (26.13)$$

Rate of change of capacitor charge · Resistance · Capacitance

The charge and current are both *exponential* functions of time. Figure 26.21a is a graph of Eq. (26.13) and Fig. 26.21b is a graph of Eq. (26.12).

Time Constant

After a time equal to RC, the current in the R-C circuit has decreased to $1/e$ (about 0.368) of its initial value. At this time, the capacitor charge has reached $(1 - 1/e) = 0.632$ of its final value $Q_f = C\mathcal{E}$. The product RC is therefore a measure of how quickly the capacitor charges. We call RC the **time constant,** or the **relaxation time,** of the circuit, denoted by τ:

$$\tau = RC \qquad \text{(time constant for } R\text{-}C \text{ circuit)} \qquad (26.14)$$

When τ is small, the capacitor charges quickly; when it is larger, the charging takes more time. If the resistance is small, it's easier for current to flow, and the capacitor charges more quickly. If R is in ohms and C in farads, τ is in seconds.

In Fig. 26.21a the horizontal axis is an *asymptote* for the curve. Strictly speaking, i never becomes exactly zero. But the longer we wait, the closer it gets. After a time equal to $10RC$, the current has decreased to 0.000045 of its initial value. Similarly, the curve in Fig. 26.21b approaches the horizontal dashed line labeled Q_f as an asymptote. The charge q never attains exactly this value, but after a time equal to $10RC$, the difference between q and Q_f is only 0.000045 of Q_f. We invite you to verify that the product RC has units of time.

Discharging a Capacitor

Now suppose that after the capacitor in Fig. 26.21b has acquired a charge Q_0, we remove the battery from our R-C circuit and connect points a and c to an open switch (**Fig. 26.22a**). We then close the switch and at the same instant reset our stopwatch to $t = 0$; at that time, $q = Q_0$. The capacitor then *discharges* through the resistor, and its charge eventually decreases to zero.

Again let i and q represent the time-varying current and charge at some instant after the connection is made. In Fig. 26.22b we make the same choice of the positive direction for current as in Fig. 26.20b. Then Kirchhoff's loop rule gives Eq. (26.10) but with $\mathcal{E} = 0$; that is,

$$i = \frac{dq}{dt} = -\frac{q}{RC} \qquad (26.15)$$

The current i is now negative; this is because positive charge q is leaving the left-hand capacitor plate in Fig. 26.22b, so the current is in the direction opposite to that shown. At time $t = 0$, when $q = Q_0$, the initial current is $I_0 = -Q_0/RC$.

To find q as a function of time, we rearrange Eq. (26.15), again change the variables to q' and t', and integrate. This time the limits for q' are Q_0 to q:

$$\int_{Q_0}^{q} \frac{dq'}{q'} = -\frac{1}{RC}\int_0^t dt'$$

$$\ln\frac{q}{Q_0} = -\frac{t}{RC}$$

***R-C* circuit, discharging capacitor:**

Capacitor charge; Initial capacitor charge; Capacitance; Time since switch closed; Resistance

$$q = Q_0 e^{-t/RC} \qquad (26.16)$$

The instantaneous current i is the derivative of this with respect to time:

***R-C* circuit, discharging capacitor:**

Current; Initial capacitor charge; Capacitance; Time since switch closed; Rate of change of capacitor charge; Resistance; Initial current $= -Q_0/RC$

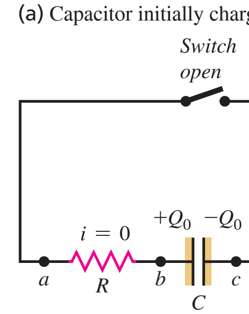
$$i = \frac{dq}{dt} = -\frac{Q_0}{RC}e^{-t/RC} = I_0 e^{-t/RC} \qquad (26.17)$$

We graph the current and the charge in **Fig. 26.23**; both quantities approach zero exponentially with time. Comparing these results with Eqs. (26.12) and (26.13), we note that the expressions for the current are identical, apart from the sign of I_0. The capacitor charge approaches zero asymptotically in Eq. (26.16), while the *difference* between q and Q approaches zero asymptotically in Eq. (26.12).

Energy considerations give us additional insight into the behavior of an R-C circuit. While the capacitor is charging, the instantaneous rate at which the battery delivers energy to the circuit is $P = \mathcal{E}i$. The instantaneous rate at which electrical energy is dissipated in

Figure **26.22** Discharging a capacitor. (a) Before the switch is closed at time $t = 0$, the capacitor charge is Q_0 and the current is zero. (b) At time t after the switch is closed, the capacitor charge is q and the current is i. The actual current direction is opposite to the direction shown; i is negative. After a long time, q and i both approach zero.

(a) Capacitor initially charged



(b) Discharging the capacitor

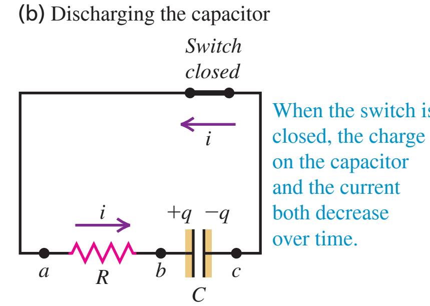
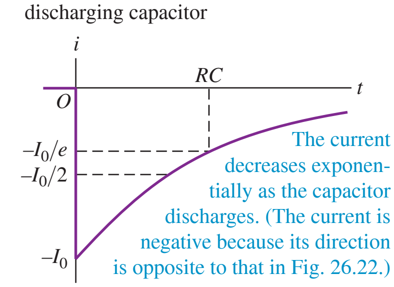
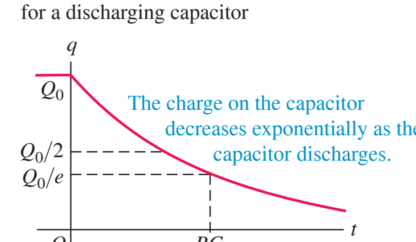


Figure **26.23** Current i and capacitor charge q as functions of time for the circuit of Fig. 26.22. The initial current is I_0 and the initial capacitor charge is Q_0. Both i and q asymptotically approach zero.

(a) Graph of current versus time for a discharging capacitor



(b) Graph of capacitor charge versus time for a discharging capacitor



the resistor is i^2R, and the rate at which energy is stored in the capacitor is $iv_{bc} = iq/C$. Multiplying Eq. (26.9) by i, we find

$$\mathcal{E}i = i^2R + \frac{iq}{C} \tag{26.18}$$

This means that of the power $\mathcal{E}i$ supplied by the battery, part (i^2R) is dissipated in the resistor and part (iq/C) is stored in the capacitor.

The *total* energy supplied by the battery during charging of the capacitor equals the battery emf $\mathcal{E}$ multiplied by the total charge Q_f, or $\mathcal{E}Q_f$. The total energy stored in the capacitor, from Eq. (24.9), is $Q_f\mathcal{E}/2$. Thus, of the energy supplied by the battery, *exactly half* is stored in the capacitor, and the other half is dissipated in the resistor. This half-and-half division of energy doesn't depend on C, R, or $\mathcal{E}$. You can verify this result by taking the integral over time of each of the power quantities in Eq. (26.18).

EXAMPLE 26.12 Charging a capacitor

WITH VARIATION PROBLEMS

A 10 MΩ resistor is connected in series with a 1.0 μF capacitor and a battery with emf 12.0 V. Before the switch is closed at time $t = 0$, the capacitor is uncharged. (a) What is the time constant? (b) What fraction of the final charge Q_f is on the capacitor at $t = 46$ s? (c) What fraction of the initial current I_0 is still flowing at $t = 46$ s?

IDENTIFY and SET UP This is the situation shown in Fig. 26.20, with $R = 10\ \text{M}\Omega$, $C = 1.0\ \mu\text{F}$, and $\mathcal{E} = 12.0\ \text{V}$. The charge q and current i vary with time as shown in Fig. 26.21. Our target variables are (a) the time constant τ, (b) the ratio q/Q_f at $t = 46$ s, and (c) the ratio i/I_0 at $t = 46$ s. Equation (26.14) gives τ. For a capacitor being charged, Eq. (26.12) gives q and Eq. (26.13) gives i.

EXECUTE (a) From Eq. (26.14),

$$\tau = RC = (10 \times 10^6\ \Omega)(1.0 \times 10^{-6}\ \text{F}) = 10\ \text{s}$$

(b) From Eq. (26.12),

$$\frac{q}{Q_f} = 1 - e^{-t/RC} = 1 - e^{-(46\ \text{s})/(10\ \text{s})} = 0.99$$

(c) From Eq. (26.13),

$$\frac{i}{I_0} = e^{-t/RC} = e^{-(46\ \text{s})/(10\ \text{s})} = 0.010$$

EVALUATE After 4.6 time constants the capacitor is 99% charged and the charging current has decreased to 1.0% of its initial value. The circuit would charge more rapidly if we reduced the time constant by using a smaller resistance.

KEYCONCEPT When a capacitor is charging in a circuit with a resistor and a source of emf, the capacitor charge q and the current i both vary with time: q approaches its final value asymptotically and i approaches zero asymptotically. The time constant for both quantities is the product RC of the resistance and the capacitance of the circuit elements.

EXAMPLE 26.13 Discharging a capacitor

WITH VARIATION PROBLEMS

The resistor and capacitor of Example 26.12 are reconnected as shown in Fig. 26.22. The capacitor has an initial charge of 5.0 μC and is discharged by closing the switch at $t = 0$. (a) At what time will the charge be 0.50 μC? (b) What is the current at this time?

IDENTIFY and SET UP Now the capacitor is being discharged, so q and i vary with time as in Fig. 26.23, with $Q_0 = 5.0 \times 10^{-6}$ C. Again we have $RC = \tau = 10$ s. Our target variables are (a) the value of t at which $q = 0.50\ \mu\text{C}$ and (b) the value of i at this time. We first solve Eq. (26.16) for t, and then solve Eq. (26.17) for i.

EXECUTE (a) Solving Eq. (26.16) for the time t gives

$$t = -RC \ln\frac{q}{Q_0} = -(10\ \text{s})\ln\frac{0.50\ \mu\text{C}}{5.0\ \mu\text{C}}$$
$$= 23\ \text{s} = 2.3\tau$$

(b) From Eq. (26.17), with $Q_0 = 5.0\ \mu\text{C} = 5.0 \times 10^{-6}$ C,

$$i = -\frac{Q_0}{RC}e^{-t/RC} = -\frac{5.0 \times 10^{-6}\ \text{C}}{10\ \text{s}}e^{-2.3} = -5.0 \times 10^{-8}\ \text{A}$$

EVALUATE The current in part (b) is negative because i has the opposite sign when the capacitor is discharging than when it is charging. Note that we could have avoided evaluating $e^{-t/RC}$ by noticing that at the time in question, $q = 0.10Q_0$; from Eq. (26.16) this means that $e^{-t/RC} = 0.10$.

KEYCONCEPT When a capacitor is discharging in a circuit with a resistor, the capacitor charge q and the current i both vary with time, and both approach zero asymptotically. The time constant for both quantities is the product RC of the resistance and the capacitance of the circuit elements.

TEST YOUR UNDERSTANDING OF SECTION 26.4 The energy stored in a capacitor is equal to $q^2/2C$. When a capacitor is discharged, what fraction of the initial energy remains after an elapsed time of one time constant? (i) $1/e$; (ii) $1/e^2$; (iii) $1 - 1/e$; (iv) $(1 - 1/e)^2$; (v) answer depends on how much energy was stored initially.

ANSWER

(ii) After one time constant, $t = RC$ and the initial charge Q_0 has decreased to $Q_0 e^{-t/RC} = Q_0 e^{-RC/RC} = Q_0 e^{-1} = Q_0/e$. Hence the stored energy has decreased from $Q_0^2/2C$ to $(Q_0/e)^2/2C = Q_0^2/2Ce^2$, a fraction $1/e^2 = 0.135$ of its initial value. This result doesn't depend on the initial value of the energy.

26.5 POWER DISTRIBUTION SYSTEMS

We conclude this chapter with a brief discussion of practical household and automotive electric-power distribution systems. Automobiles use direct-current (dc) systems, while nearly all household, commercial, and industrial systems use alternating current (ac) because of the ease of stepping voltage up and down with transformers. Most of the same basic wiring concepts apply to both. We'll talk about alternating-current circuits in greater detail in Chapter 31.

The various lamps, motors, and other appliances to be operated are always connected in *parallel* to the power source (the wires from the power company for houses, or from the battery and alternator for a car). If appliances were connected in series, shutting one appliance off would shut them all off (see Example 26.2 in Section 26.1). **Figure 26.24** shows the basic idea of house wiring. One side of the "line," as the pair of conductors is called, is called the *neutral* side; it is always connected to "ground" at the entrance panel. For houses, *ground* is an actual electrode driven into the earth (which is usually a good conductor) or sometimes connected to the household water pipes. Electricians speak of the "hot" side and the "neutral" side of the line. Most modern house wiring systems have *two* hot lines with opposite polarity with respect to the neutral. We'll return to this detail later.

Household voltage is nominally 120 V in the United States and Canada, and often 240 V in Europe. (For alternating current, which varies sinusoidally with time, these numbers represent the *root-mean-square* voltage, which is $1/\sqrt{2}$ times the peak voltage. We'll discuss this further in Section 31.1.) The amount of current I drawn by a given device is determined by its power input P, given by Eq. (25.17): $P = VI$. Hence $I = P/V$. For example, the current in a 100 W light bulb is

$$I = \frac{P}{V} = \frac{100\ \text{W}}{120\ \text{V}} = 0.83\ \text{A}$$

The power input to this bulb is actually determined by its resistance R. Using Eq. (25.18), which states that $P = VI = I^2R = V^2/R$ for a resistor, we get the resistance of this bulb at operating temperature:

$$R = \frac{V}{I} = \frac{120\ \text{V}}{0.83\ \text{A}} = 144\ \Omega \qquad \text{or} \qquad R = \frac{V^2}{P} = \frac{(120\ \text{V})^2}{100\ \text{W}} = 144\ \Omega$$

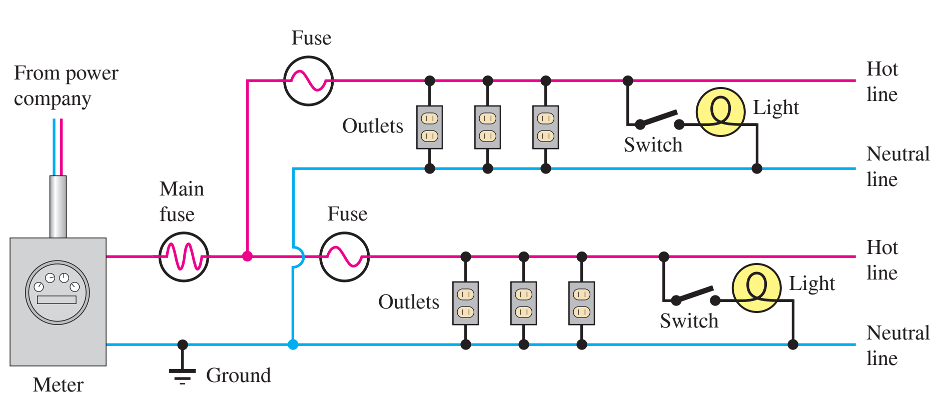


Figure **26.24** Schematic diagram of part of a house wiring system. Only two branch circuits are shown; an actual system might have four to thirty branch circuits. Lamps and appliances may be plugged into the outlets. The grounding wires, which normally carry no current, are not shown.

Similarly, a 1500 W waffle iron draws a current of $(1500\ \text{W})/(120\ \text{V}) = 12.5\ \text{A}$ and has a resistance, at operating temperature, of $9.6\ \Omega$. Because of the temperature dependence of resistivity, the resistances of these devices are considerably less when they are cold. If you measure the resistance of a 100 W light bulb with an ohmmeter (whose small current causes very little temperature rise), you'll probably get a value of about $10\ \Omega$. When a light bulb is turned on, this low resistance causes an initial surge of current until the filament heats up. That's why a light bulb that's ready to burn out nearly always does so just when you turn it on.

Circuit Overloads and Short Circuits

The maximum current available from an individual circuit is limited by the resistance of the wires. As we discussed in Section 25.5, the I^2R power loss in the wires causes them to become hot, and in extreme cases this can cause a fire or melt the wires. Ordinary lighting and outlet wiring in houses usually uses 12 gauge wire. This has a diameter of 2.05 mm and can carry a maximum current of 20 A safely (without overheating). Larger-diameter wires of the same length have lower resistance [see Eq. (25.10)]. Hence 8 gauge (3.26 mm) or 6 gauge (4.11 mm) is used for high-current appliances such as clothes dryers, and 2 gauge (6.54 mm) or larger is used for the main power lines entering a house.

Figure **26.25** (a) Excess current will melt the thin wire of lead–tin alloy that runs along the length of a fuse, inside the transparent housing. (b) The switch on this circuit breaker will flip if the maximum allowable current is exceeded.

(a)

(b)

Protection against overloading and overheating of circuits is provided by fuses or circuit breakers. A *fuse* contains a link of lead–tin alloy with a very low melting temperature; the link melts and breaks the circuit when its rated current is exceeded (**Fig. 26.25a**). A *circuit breaker* is an electromechanical device that performs the same function, using an electromagnet or a bimetallic strip to "trip" the breaker and interrupt the circuit when the current exceeds a specified value (Fig. 26.25b). Circuit breakers have the advantage that they can be reset after they are tripped, while a blown fuse must be replaced.

CAUTION Fuses If your system has fuses and you plug too many high-current appliances into the same outlet, the fuse blows. Do not replace the fuse with one that has a higher rating; if you do, you risk overheating the wires and starting a fire. The only safe solution is to distribute the appliances among several circuits. Modern kitchens often have three or four separate 20 A circuits.

Contact between the hot and neutral sides of the line causes a *short circuit*. Such a situation, which can be caused by faulty insulation or by a variety of mechanical malfunctions, provides a very low-resistance current path, permitting a very large current that would quickly melt the wires and ignite their insulation if the current were not interrupted by a fuse or circuit breaker (see Example 25.10 in Section 25.5). An equally dangerous situation is a broken wire that interrupts the current path, creating an *open circuit*. This is hazardous because of the sparking that can occur at the point of intermittent contact.

In approved wiring practice, a fuse or breaker is placed *only* in the hot side of the line, never in the neutral side. Otherwise, if a short circuit should develop because of faulty insulation or other malfunction, the ground-side fuse could blow. The hot side would still be live and would pose a shock hazard if you touched the live conductor and a grounded object such as a water pipe. For similar reasons the wall switch for a light fixture should always be in the hot side of the line, never the neutral side.

Further protection against shock hazard is provided by a third conductor called the *grounding wire,* included in all present-day wiring. This conductor corresponds to the long round or U-shaped prong of the three-prong connector plug on an appliance or power tool. It is connected to the neutral side of the line at the entrance panel. The grounding wire normally carries no current, but it connects the metal case or frame of the device to ground. If a conductor on the hot side of the line accidentally contacts the frame or case, the grounding conductor provides a current path, and the fuse blows. Without the ground wire, the frame could become "live"—that is, at a potential 120 V above ground. Then if you touched it and a water pipe (or even a damp floor) at the same time, you could get a

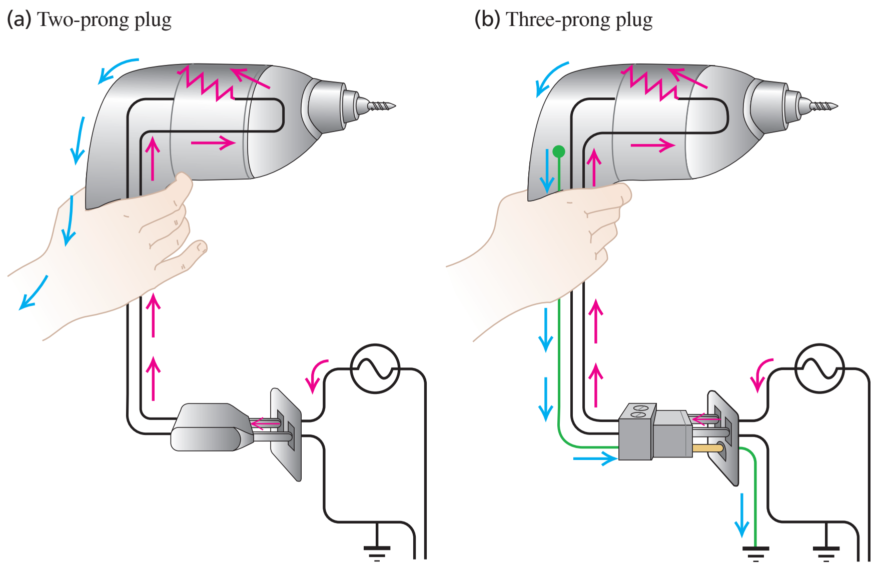


Figure **26.26** (a) If a malfunctioning electric drill is connected to a wall socket via a two-prong plug, a person may receive a shock. (b) When the drill malfunctions when connected via a three-prong plug, a person touching it receives no shock, because electric charge flows through the ground wire (shown in green) to the third prong and into the ground rather than into the person's body. If the ground current is appreciable, the fuse blows.

dangerous shock (**Fig. 26.26**). In some situations, especially outlets located outdoors or near a sink or other water pipes, a special kind of circuit breaker called a *ground-fault interrupter* (GFI or GFCI) is used. This device senses the difference in current between the hot and neutral conductors (which is normally zero) and trips when some very small value, typically 5 mA, is exceeded.

Household and Automotive Wiring

Most modern household wiring systems actually use a slight elaboration of the system described above. The power company provides *three* conductors. One is neutral; the other two are both at 120 V with respect to the neutral but with opposite polarity, giving a voltage between them of 240 V. The power company calls this a *three-wire line,* in contrast to the 120 V two-wire (plus ground wire) line described above. With a three-wire line, 120 V lamps and appliances can be connected between neutral and either hot conductor, and high-power devices requiring 240 V, such as electric ranges and clothes dryers, are connected between the two hot lines.

All of the above discussion can be applied directly to automobile wiring. The voltage is about 13 V (direct current); the power is supplied by the battery and by the alternator, which charges the battery when the engine is running. The neutral side of each circuit is connected to the body and frame of the vehicle. For this low voltage a separate grounding conductor is not required for safety. The fuse or circuit breaker arrangement is the same in principle as in household wiring. Because of the lower voltage (less energy per charge), more current (a greater number of charges per second) is required for the same power; a 100 W headlight bulb requires a current of about $(100\text{ W})/(13\text{ V}) = 8\text{ A}$.

Although we spoke of *power* in the above discussion, what we buy from the power company is *energy*. Power is energy transferred per unit time, so energy is average power multiplied by time. The usual unit of energy sold by the power company is the kilowatt-hour $(1\text{ kW}\cdot\text{h})$:

$$1\text{ kW}\cdot\text{h} = (10^3\text{ W})(3600\text{ s}) = 3.6\times10^6\text{ W}\cdot\text{s} = 3.6\times10^6\text{ J}$$

In the United States, one kilowatt-hour typically costs 8 to 27 cents, depending on the location and quantity of energy purchased. To operate a 1500 W (1.5 kW) waffle iron continuously for 1 hour requires $1.5\text{ kW}\cdot\text{h}$ of energy; at 10 cents per kilowatt-hour, the energy cost is 15 cents. The cost of operating any lamp or appliance for a specified time can be calculated in the same way if the power rating is known. However, many electric cooking utensils (including waffle irons) cycle on and off to maintain a constant temperature, so the average power may be less than the power rating marked on the device.

EXAMPLE 26.14 A kitchen circuit

An 1800 W toaster, a 1.3 kW electric frying pan, and a 100 W lamp are plugged into the same 20 A, 120 V circuit. (a) What current is drawn by each device, and what is the resistance of each device? (b) Will this combination trip the circuit breaker?

IDENTIFY and SET UP When plugged into the same circuit, the three devices are connected in parallel, so the voltage across each appliance is $V = 120$ V. We find the current I drawn by each device from the relationship $P = VI$, where P is the power input of the device. To find the resistance R of each device we use the relationship $P = V^2/R$.

EXECUTE (a) To simplify the calculation of current and resistance, we note that $I = P/V$ and $R = V^2/P$. Hence

$$I_{\text{toaster}} = \frac{1800\text{ W}}{120\text{ V}} = 15\text{ A} \qquad R_{\text{toaster}} = \frac{(120\text{ V})^2}{1800\text{ W}} = 8\ \Omega$$

$$I_{\text{frying pan}} = \frac{1300\text{ W}}{120\text{ V}} = 11\text{ A} \qquad R_{\text{frying pan}} = \frac{(120\text{ V})^2}{1300\text{ W}} = 11\ \Omega$$

$$I_{\text{lamp}} = \frac{100\text{ W}}{120\text{ V}} = 0.83\text{ A} \qquad R_{\text{lamp}} = \frac{(120\text{ V})^2}{100\text{ W}} = 144\ \Omega$$

For constant voltage the device with the *least* resistance (in this case the toaster) draws the most current and receives the most power.

(b) The total current through the line is the sum of the currents drawn by the three devices:

$$\begin{aligned} I &= I_{\text{toaster}} + I_{\text{frying pan}} + I_{\text{lamp}} \\ &= 15\text{ A} + 11\text{ A} + 0.83\text{ A} = 27\text{ A} \end{aligned}$$

This exceeds the 20 A rating of the line, and the circuit breaker will indeed trip.

EVALUATE We could also find the total current by using $I = P/V$ and the total power P delivered to all three devices:

$$\begin{aligned} I &= \frac{P_{\text{toaster}} + P_{\text{frying pan}} + P_{\text{lamp}}}{V} \\ &= \frac{1800\text{ W} + 1300\text{ W} + 100\text{ W}}{120\text{ V}} = 27\text{ A} \end{aligned}$$

A third way to determine I is to use $I = V/R_{\text{eq}}$, where R_{eq} is the equivalent resistance of the three devices in parallel:

$$I = \frac{V}{R_{\text{eq}}} = (120\text{ V})\left(\frac{1}{8\ \Omega} + \frac{1}{11\ \Omega} + \frac{1}{144\ \Omega}\right) = 27\text{ A}$$

Appliances with such current demands are common, so modern kitchens have more than one 20 A circuit. To keep currents safely below 20 A, the toaster and frying pan should be plugged into different circuits.

KEYCONCEPT The voltage across any device plugged into a household wiring system has the same value. The smaller the resistance of the device, the more current it draws and the more power it receives.

TEST YOUR UNDERSTANDING OF SECTION 26.5 To prevent the circuit breaker in Example 26.14 from blowing, a home electrician replaces the circuit breaker with one rated at 40 A. Is this a reasonable thing to do?

ANSWER

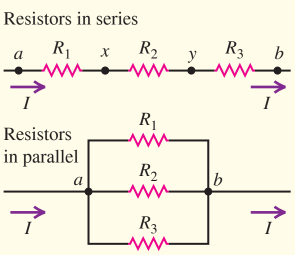
no This is a very dangerous thing to do. The circuit breaker will allow currents up to 40 A, double the rated value of the wiring. The amount of power $P = I^2R$ dissipated in a section of wire can therefore be up to four times the rated value, so the wires could get hot and start a fire. (This assumes the resistance R remains unchanged. In fact, R increases with temperature, so the dissipated power can be even greater, and more dangerous, than we have estimated.)

CHAPTER 26 SUMMARY

Resistors in series and parallel: When several resistors $R_1, R_2, R_3, \ldots$ are connected in series, the equivalent resistance R_{eq} is the sum of the individual resistances. The same *current* flows through all the resistors in a series connection. When several resistors are connected in parallel, the reciprocal of equivalent resistance R_{eq} is the sum of the reciprocals of the individual resistances. All resistors in a parallel connection have the same *potential difference* between their terminals. (See Examples 26.1 and 26.2.)

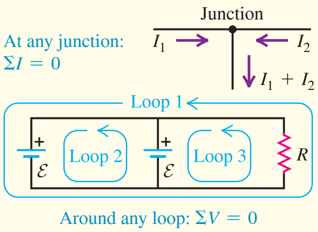
$$R_{eq} = R_1 + R_2 + R_3 + \cdots \tag{26.1}$$
(resistors in series)

$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots \tag{26.2}$$
(resistors in parallel)

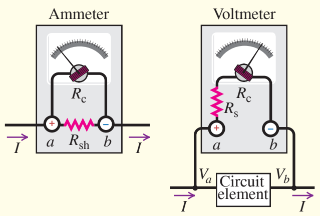


Kirchhoff's rules: Kirchhoff's junction rule is based on conservation of charge. It states that the algebraic sum of the currents into any junction must be zero. Kirchhoff's loop rule is based on conservation of energy and the conservative nature of electrostatic fields. It states that the algebraic sum of potential differences around any loop must be zero. Careful use of consistent sign rules is essential in applying Kirchhoff's rules. (See Examples 26.3–26.7.)

$$\sum I = 0 \quad \text{(junction rule)} \tag{26.5}$$
$$\sum V = 0 \quad \text{(loop rule)} \tag{26.6}$$



Electrical measuring instruments: In a d'Arsonval galvanometer, the deflection is proportional to the current in the coil. For a larger current range, a shunt resistor is added, so some of the current bypasses the meter coil. Such an instrument is called an ammeter. If the coil and any additional series resistance included obey Ohm's law, the meter can also be calibrated to read potential difference or voltage. The instrument is then called a voltmeter. A good ammeter has very low resistance; a good voltmeter has very high resistance. (See Examples 26.8–26.11.)



***R-C* circuits:** When a capacitor is charged by a battery in series with a resistor, the current and capacitor charge are not constant. The charge approaches its final value asymptotically and the current approaches zero asymptotically. The charge and current in the circuit are given by Eqs. (26.12) and (26.13). After a time $\tau = RC$, the charge has approached within $1/e$ of its final value. This time is called the time constant or relaxation time of the circuit. When the capacitor discharges, the charge and current are given as functions of time by Eqs. (26.16) and (26.17). The time constant is the same for charging and discharging. (See Examples 26.12 and 26.13.)

Capacitor charging:

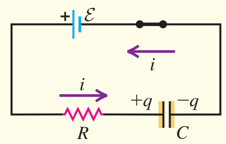
$$\begin{aligned} q &= C\mathcal{E}\left(1 - e^{-t/RC}\right) \\ &= Q_f\left(1 - e^{-t/RC}\right) \end{aligned} \tag{26.12}$$

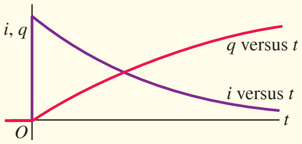
$$i = \frac{dq}{dt} = \frac{\mathcal{E}}{R}e^{-t/RC} = I_0 e^{-t/RC} \tag{26.13}$$

Capacitor discharging:

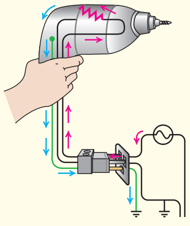
$$q = Q_0 e^{-t/RC} \tag{26.16}$$

$$i = \frac{dq}{dt} = -\frac{Q_0}{RC}e^{-t/RC} = I_0 e^{-t/RC} \tag{26.17}$$





Household wiring: In household wiring systems, the various electrical devices are connected in parallel across the power line, which consists of a pair of conductors, one "hot" and the other "neutral." An additional "ground" wire is included for safety. The maximum permissible current in a circuit is determined by the size of the wires and the maximum temperature they can tolerate. Protection against excessive current and the resulting fire hazard is provided by fuses or circuit breakers. (See Example 26.14.)



Chapter 26 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 26.1 and 26.2 (Section 26.1) before attempting these problems.

VP26.2.1 You have three resistors: $R_1 = 1.00\ \Omega$, $R_2 = 2.00\ \Omega$, and $R_3 = 4.00\ \Omega$. Find the equivalent resistance for the combinations shown in (a) Fig. 26.1a, (b) Fig. 26.1b, (c) Fig. 26.1c, and (d) Fig. 26.1d.

VP26.2.2 You have three resistors, $R_1 = 4.00\ \Omega$, $R_2 = 5.00\ \Omega$, and $R_3 = 6.00\ \Omega$, connected as shown in Fig. 26.1d. The terminals of a battery with emf 24.0 V and negligible internal resistance are connected to points a and b. Find the current through (a) the battery, (b) R_1, (c) R_2, and (d) R_3.

VP26.2.3 You have three resistors, $R_1 = 7.00\ \Omega$, $R_2 = 8.00\ \Omega$, and $R_3 = 9.00\ \Omega$, connected as shown in Fig. 26.1b. The terminals of a battery with emf 12.0 V and negligible internal resistance are connected to points a and b. Find (a) the power output of the battery and the power input to (b) R_1, (c) R_2, and (d) R_3.

VP26.2.4 You have three resistors, $R_1 = 5.00\ \Omega$, $R_2 = 6.00\ \Omega$, and $R_3 = 7.00\ \Omega$, connected as shown in Fig. 26.1c. The terminals of a battery with emf 9.00 V and negligible internal resistance are connected to points a and b. Find (a) the power output of the battery and the power input to (b) R_1, (c) R_2, and (d) R_3.

Be sure to review EXAMPLES 26.3, 26.4, 26.5, 26.6, and 26.7 (Section 26.2) before attempting these problems.

VP26.7.1 In the circuit shown in Fig. 26.10a, you reverse the 4 V battery so that its positive terminal is on the right side instead of the left. Find (a) the current in the circuit, (b) the potential difference V_{ab}, and (c) the power output of the emf of each battery.

VP26.7.2 Figure 26.6a shows two batteries, one with emf $\mathcal{E}_1$ and one with emf $\mathcal{E}_2$, connected to a resistance R. Current I_1 flows through emf $\mathcal{E}_1$ from point c toward point a, and current I_2 flows through emf $\mathcal{E}_2$ from point b toward point a. If $\mathcal{E}_1 = 8.00$ V, $\mathcal{E}_2 = 9.00$ V, $R = 5.00\ \Omega$, $I_1 = 0.200$ A, and $I_2 = 1.35$ A, find the values of (a) the potential difference V_{ab}, (b) the internal resistance r_1, and (c) the internal resistance r_2.

VP26.7.3 In the circuit shown in Figure 26.12, you replace the 2 Ω resistor between points b and d with a 3 Ω resistor. Find the power dissipated in (a) the 1 Ω resistor between points c and a, (b) the 1 Ω resistor between points c and b, (c) the 1 Ω resistor between points a and b, (d) the 1 Ω resistor between points a and d, and (e) the 3 Ω resistor between points b and d. Give each answer to three significant figures.

VP26.7.4 In the circuit shown in Fig. 26.9a, let $r_1 = r_2 = 0$ (that is, the internal resistances are very small) and let $R_1 = R_2 = R_3 = R$. Find the values of (a) I_1, (b) I_2, and (c) I_3.

Be sure to review EXAMPLES 26.12 and 26.13 (Section 26.4) before attempting these problems.

VP26.13.1 You connect a 10.0 MΩ resistor in series with a 3.20 μF capacitor and a battery with emf 9.00 V. Before you close the switch at $t = 0$ to complete the circuit, the capacitor is uncharged. Find (a) the final capacitor charge, (b) the initial current, (c) the time constant, (d) the fraction of the final charge on the capacitor at $t = 18.0$ s, and (e) the fraction of the initial current present at $t = 18.0$ s.

VP26.13.2 A 2.20 μF capacitor initially has charge 4.20 μC. You connect it in series with a 4.00 MΩ resistor and an open switch. If $t = 0$ is the time when you close the switch, (a) at what time will the capacitor charge be 1.20 μC and (b) what will be the current in the circuit at this time?

VP26.13.3 You connect an unknown resistor in series with an 8.00 μF capacitor and an open switch. Before you close the switch at $t = 0$, the capacitor has charge 5.50 μC; at $t = 17.0$ s, the charge has decreased to 1.10 μC. Find (a) the resistance of the resistor, (b) the current in the circuit just after you close the switch, and (c) the current in the circuit at $t = 17.0$ s.

VP26.13.4 You connect an initially uncharged 6.40 μF capacitor in series with a 5.00 MΩ resistor and a battery with emf 12.0 V. After letting the capacitor charge for 51.0 s, you disconnect it from this circuit and connect it in series to an open switch and a 6.00 MΩ resistor. Find the charge on the capacitor (a) when you disconnect it from the first circuit and (b) 70.0 s after you close the switch in the second circuit.

BRIDGING PROBLEM Two Capacitors and Two Resistors

A 2.40 μF capacitor and a 3.60 μF capacitor are connected in series. (a) A charge of 5.20 mC is placed on each capacitor. What is the energy stored in the capacitors? (b) A 655 Ω resistor is connected to the terminals of the capacitor combination, and a voltmeter with resistance $4.58 \times 10^4\ \Omega$ is connected across the resistor (**Fig. 26.27**). What is the rate of change of the energy stored in the capacitors just after the connection is made? (c) How long after the connection is made has the energy stored in the capacitors decreased to $1/e$ of its initial value? (d) At the instant calculated in part (c), what is the rate of change of the energy stored in the capacitors?

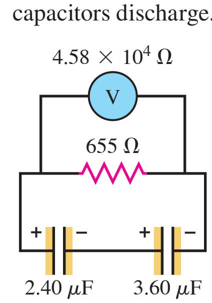


Figure 26.27 When the connection is made, the charged capacitors discharge.

SOLUTION GUIDE

IDENTIFY and SET UP

1. The two capacitors act as a single equivalent capacitor (see Section 24.2), and the resistor and voltmeter act as a single equivalent resistor. Select equations that will allow you to calculate the values of these equivalent circuit elements.
2. In part (a) you'll need to use Eq. (24.9), which gives the energy stored in a capacitor.
3. For parts (b), (c), and (d), you'll need to use Eq. (24.9) as well as Eqs. (26.16) and (26.17), which give the capacitor charge and current as functions of time. (*Hint:* The rate at which energy is lost by the capacitors equals the rate at which energy is dissipated in the resistances.)

EXECUTE

4. Find the stored energy at $t = 0$.
5. Find the rate of change of the stored energy at $t = 0$.
6. Find the value of t at which the stored energy has $1/e$ of the value you found in step 4.
7. Find the rate of change of the stored energy at the time you found in step 6.

EVALUATE

8. Check your results from steps 5 and 7 by calculating the rate of change in a different way. (*Hint:* The rate of change of the stored energy U is dU/dt.)

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

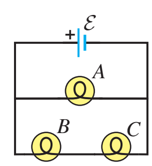
Q26.1 In which 120 V light bulb does the filament have greater resistance: a 60 W bulb or a 120 W bulb? If the two bulbs are connected to a 120 V line in series, through which bulb will there be the greater voltage drop? What if they are connected in parallel? Explain your reasoning.

Q26.2 Two 120 V light bulbs, one 25 W and one 200 W, were connected in series across a 240 V line. It seemed like a good idea at the time, but one bulb burned out almost immediately. Which one burned out, and why?

Q26.3 You connect a number of identical light bulbs to a flashlight battery. (a) What happens to the brightness of each bulb as more and more bulbs are added to the circuit if you connect them (i) in series and (ii) in parallel? (b) Will the battery last longer if the bulbs are in series or in parallel? Explain your reasoning.

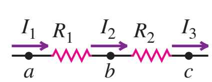
Q26.4 In the circuit shown in **Fig. Q26.4**, three identical light bulbs are connected to a flashlight battery. How do the brightnesses of the bulbs compare? Which light bulb has the greatest current passing through it? Which light bulb has the greatest potential difference between its terminals? What happens if bulb A is unscrewed? Bulb B? Bulb C? Explain your reasoning.

Figure **Q26.4**



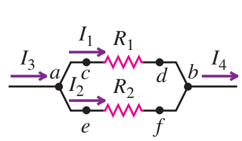
Q26.5 If two resistors R_1 and R_2 $(R_2 > R_1)$ are connected in series as shown in **Fig. Q26.5**, which of the following must be true? In each case justify your answer. (a) $I_1 = I_2 = I_3$. (b) The current is greater in R_1 than in R_2. (c) The electrical power consumption is the same for both resistors. (d) The electrical power consumption is greater in R_2 than in R_1. (e) The potential drop is the same across both resistors. (f) The potential at point a is the same as at point c. (g) The potential at point b is lower than at point c. (h) The potential at point c is lower than at point b.

Figure **Q26.5**



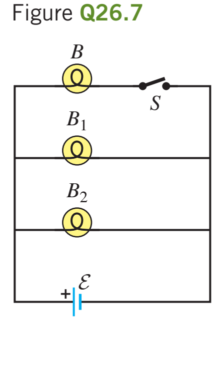
Q26.6 If two resistors R_1 and R_2 $(R_2 > R_1)$ are connected in parallel as shown in **Fig. Q26.6**, which of the following must be true? In each case justify your answer. (a) $I_1 = I_2$. (b) $I_3 = I_4$. (c) The current is greater in R_1 than in R_2. (d) The rate of electrical energy consumption is the same for both resistors. (e) The rate of electrical energy consumption is greater in R_2 than in R_1. (f) $V_{cd} = V_{ef} = V_{ab}$. (g) Point c is at higher potential than point d. (h) Point f is at higher potential than point e. (i) Point c is at higher potential than point e.

Figure **Q26.6**



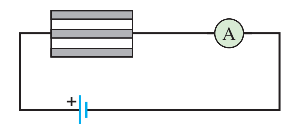
Q26.7 A battery with no internal resistance is connected across identical light bulbs as shown in **Fig. Q26.7**. When you close the switch S, will the brightness of bulbs B_1 and B_2 change? If so, how will it change? Explain.

Figure **Q26.7**



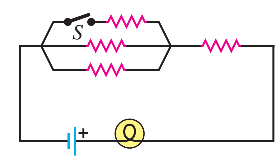
Q26.8 A resistor consists of three identical metal strips connected as shown in **Fig. Q26.8**. If one of the strips is cut out, does the ammeter reading increase, decrease, or stay the same? Why?

Figure **Q26.8**



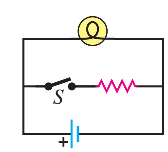
Q26.9 A light bulb is connected in the circuit shown in **Fig. Q26.9**. If we close the switch S, does the bulb's brightness increase, decrease, or remain the same? Explain why.

Figure **Q26.9**



Q26.10 A real battery, having nonnegligible internal resistance, is connected across a light bulb as shown in **Fig. Q26.10**. When the switch S is closed, what happens to the brightness of the bulb? Why?

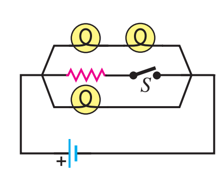
Figure **Q26.10**



Q26.11 If the battery in Discussion Question Q26.10 is ideal with no internal resistance, what will happen to the brightness of the bulb when S is closed? Why?

Q26.12 Consider the circuit shown in **Fig. Q26.12**. What happens to the brightnesses of the bulbs when the switch S is closed if the battery (a) has no internal resistance and (b) has nonnegligible internal resistance? Explain why.

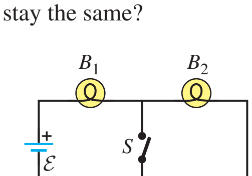
Figure **Q26.12**



Q26.13 Is it possible to connect resistors together in a way that cannot be reduced to some combination of series and parallel combinations? If so, give examples. If not, state why not.

Q26.14 The battery in the circuit shown in **Fig. Q26.14** has no internal resistance. After you close the switch S, will the brightness of bulb B_1 increase, decrease, or stay the same?

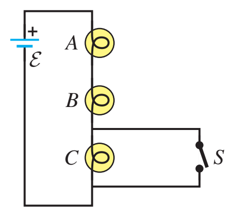
Figure **Q26.14**



Q26.15 In a two-cell flashlight, the batteries are usually connected in series. Why not connect them in parallel? What possible advantage could there be in connecting several identical batteries in parallel?

Q26.16 Identical light bulbs A, B, and C are connected as shown in **Fig. Q26.16**. When the switch S is closed, bulb C goes out. Explain why. What happens to the brightness of bulbs A and B? Explain.

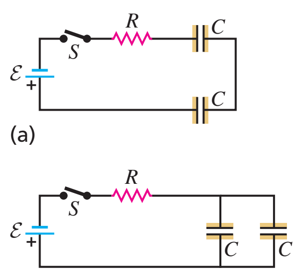
Figure **Q26.16**



Q26.17 The emf of a flashlight battery is roughly constant with time, but its internal resistance increases with age and use. What sort of meter should be used to test the freshness of a battery?

Q26.18 Will the capacitors in the circuits shown in **Fig. Q26.18** charge at the same rate when the switch S is closed? If not, in which circuit will the capacitors charge more rapidly? Explain.

Figure **Q26.18**



Q26.19 Verify that the time constant RC has units of time.

Q26.20 For very large resistances it is easy to construct R-C circuits that have time constants of several seconds or minutes. How might this fact be used to measure very large resistances, those that are too large to measure by more conventional means?

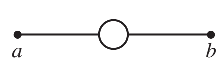
Q26.21 When a capacitor, battery, and resistor are connected in series, does the resistor affect the maximum charge stored on the capacitor? Why or why not? What purpose does the resistor serve?

EXERCISES

Section 26.1 Resistors in Series and Parallel

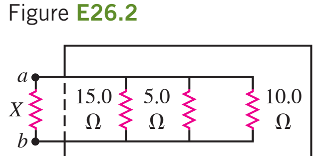
26.1 •• A uniform wire of resistance R is cut into three equal lengths. One of these is formed into a circle and connected between the other two (**Fig. E26.1**). What is the resistance between the opposite ends a and b?

Figure **E26.1**



26.2 •• A machine part has a resistor X protruding from an opening in the side. This resistor is connected to three other resistors, as shown in **Fig. E26.2**. An ohmmeter connected across a and b reads 2.00 Ω. What is the resistance of X?

Figure **E26.2**

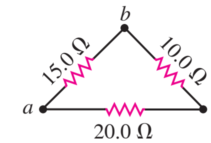


26.3 • A resistor with $R_1 = 25.0\ \Omega$ is connected to a battery that has negligible internal resistance and electrical energy is dissipated by R_1 at a rate of 36.0 W. If a second resistor with $R_2 = 15.0\ \Omega$ is connected in series with R_1, what is the total rate at which electrical energy is dissipated by the two resistors?

26.4 • A 42 Ω resistor and a 20 Ω resistor are connected in parallel, and the combination is connected across a 240 V dc line. (a) What is the resistance of the parallel combination? (b) What is the total current through the parallel combination? (c) What is the current through each resistor?

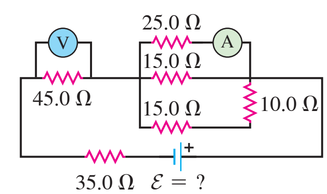
26.5 • A triangular array of resistors is shown in **Fig. E26.5**. What current will this array draw from a 35.0 V battery having negligible internal resistance if we connect it across (a) ab; (b) bc; (c) ac? (d) If the battery has an internal resistance of 3.00 Ω, what current will the array draw if the battery is connected across bc?

Figure **E26.5**



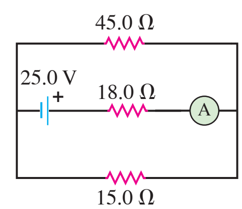
26.6 •• For the circuit shown in **Fig. E26.6** both meters are idealized, the battery has no appreciable internal resistance, and the ammeter reads 1.25 A. (a) What does the voltmeter read? (b) What is the emf $\mathcal{E}$ of the battery?

Figure **E26.6**



26.7 •• For the circuit shown in **Fig. E26.7** find the reading of the idealized ammeter if the battery has an internal resistance of 3.26 Ω.

Figure **E26.7**



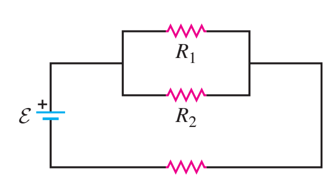
26.8 • Two resistors, R_1 and R_2, are connected in parallel to a power supply that has voltage V and negligible internal resistance. $R_2 = 8.00\ \Omega$ and the resistance of R_1 is not known. For several values of V, you measure the current I flowing through the voltage source. You plot the data as I versus V and find that they lie close to a straight line that has slope $0.208\ \Omega^{-1}$. What is the resistance of R_1?

26.9 • Six identical resistors, each with resistance R, are connected to an emf $\mathcal{E}$. (a) In terms of $\mathcal{E}$ and R, what is the current I through each of the resistors if they are connected in parallel? (b) In series? (c) For which network of resistors, series or parallel, is the power consumed in each resistor greater?

26.10 •• Power Rating of a Resistor. The *power rating* of a resistor is the maximum power the resistor can safely dissipate without too great a rise in temperature and hence damage to the resistor. (a) If the power rating of a 15 kΩ resistor is 5.0 W, what is the maximum allowable potential difference across the terminals of the resistor? (b) A 9.0 kΩ resistor is to be connected across a 120 V potential difference. What power rating is required? (c) A 100.0 Ω and a 150.0 Ω resistor, both rated at 2.00 W, are connected in series across a variable potential difference. What is the greatest this potential difference can be without overheating either resistor, and what is the rate of heat generated in each resistor under these conditions?

26.11 • In **Fig. E26.11**, $R_1 = 3.00\ \Omega$, $R_2 = 6.00\ \Omega$, and $R_3 = 5.00\ \Omega$. The battery has negligible internal resistance. The current I_2 through R_2 is 4.00 A. (a) What are the currents I_1 and I_3? (b) What is the emf of the battery?

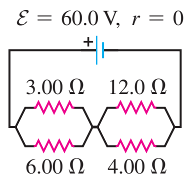
Figure **E26.11**



26.12 •• In Fig. E26.11 the battery has emf 35.0 V and negligible internal resistance. $R_1 = 5.00\ \Omega$. The current through R_1 is 1.50 A, and the current through $R_3 = 4.50$ A. What are the resistances R_2 and R_3?

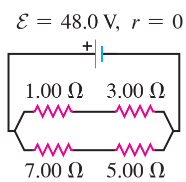
26.13 • Compute the equivalent resistance of the network in **Fig. E26.13**, and find the current in each resistor. The battery has negligible internal resistance.

Figure **E26.13**



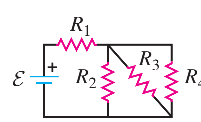
26.14 • Compute the equivalent resistance of the network in **Fig. E26.14**, and find the current in each resistor. The battery has negligible internal resistance.

Figure **E26.14**



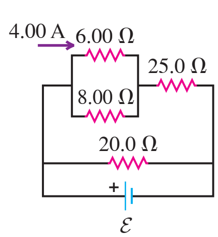
26.15 • In the circuit of **Fig. E26.15**, each resistor represents a light bulb. Let $R_1 = R_2 = R_3 = R_4 = 4.50\ \Omega$ and $\mathcal{E} = 9.00$ V. (a) Find the current in each bulb. (b) Find the power dissipated in each bulb. Which bulb or bulbs glow the brightest? (c) Bulb R_4 is now removed from the circuit, leaving a break in the wire at its position. Now what is the current in each of the remaining bulbs R_1, R_2, and R_3? (d) With bulb R_4 removed, what is the power dissipated in each of the remaining bulbs? (e) Which light bulb(s) glow brighter as a result of removing R_4? Which bulb(s) glow less brightly? Discuss why there are different effects on different bulbs.

Figure **E26.15**



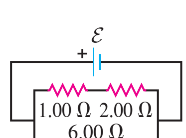
26.16 • Consider the circuit shown in **Fig. E26.16**. The current through the 6.00 Ω resistor is 4.00 A, in the direction shown. What are the currents through the 25.0 Ω and 20.0 Ω resistors?

Figure **E26.16**



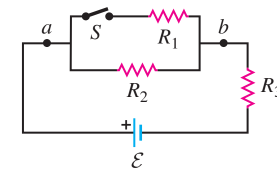
26.17 • In the circuit shown in **Fig. E26.17**, the voltage across the 2.00 Ω resistor is 12.0 V. What are the emf of the battery and the current through the 6.00 Ω resistor?

Figure **E26.17**



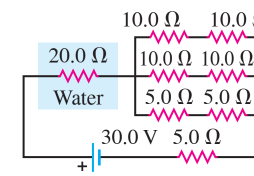
26.18 •• In the circuit shown in **Fig. E26.18**, $\mathcal{E} = 36.0$ V, $R_1 = 4.00\ \Omega$, $R_2 = 6.00\ \Omega$, and $R_3 = 3.00\ \Omega$. (a) What is the potential difference V_{ab} between points a and b when the switch S is open and when S is closed? (b) For each resistor, calculate the current through the resistor with S open and with S closed. For each resistor, does the current increase or decrease when S is closed?

Figure **E26.18**



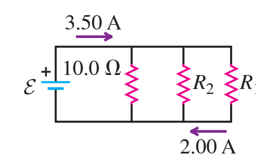
26.19 •• **CP** In the circuit in **Fig. E26.19**, a 20.0 Ω resistor is inside 100 g of pure water that is surrounded by insulating Styrofoam. If the water is initially at 10.0°C, how long will it take for its temperature to rise to 58.0°C?

Figure **E26.19**



26.20 • In the circuit shown in **Fig. E26.20**, the rate at which R_1 is dissipating electrical energy is 15.0 W. (a) Find R_1 and R_2. (b) What is the emf of the battery? (c) Find the current through both R_2 and the 10.0 Ω resistor. (d) Calculate the total electrical power consumption in all the resistors and the electrical power delivered by the battery. Show that your results are consistent with conservation of energy.

Figure **E26.20**

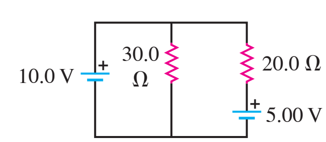


26.21 • **Light Bulbs in Series and in Parallel.** Two light bulbs have constant resistances of 400 Ω and 800 Ω. If the two light bulbs are connected in series across a 120 V line, find (a) the current through each bulb; (b) the power dissipated in each bulb; (c) the total power dissipated in both bulbs. The two light bulbs are now connected in parallel across the 120 V line. Find (d) the current through each bulb; (e) the power dissipated in each bulb; (f) the total power dissipated in both bulbs. (g) In each situation, which of the two bulbs glows the brightest? (h) In which situation is there a greater total light output from both bulbs combined?

Section 26.2 Kirchhoff's Rules

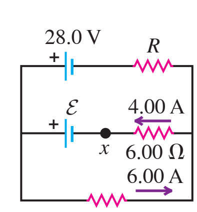
26.22 •• The batteries shown in the circuit in **Fig. E26.22** have negligibly small internal resistances. Find the current through (a) the 30.0 Ω resistor; (b) the 20.0 Ω resistor; (c) the 10.0 V battery.

Figure **E26.22**



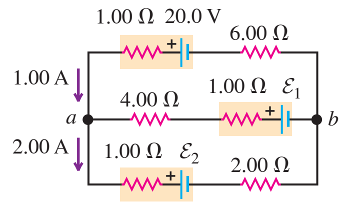
26.23 • In the circuit shown in **Fig. E26.23** find (a) the current in resistor R; (b) the resistance R; (c) the unknown emf $\mathcal{E}$. (d) If the circuit is broken at point x, what is the current in resistor R?

Figure **E26.23**



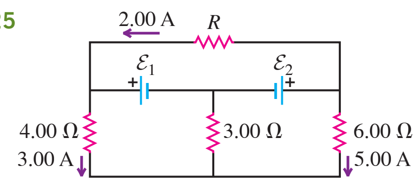
26.24 • Find the emfs $\mathcal{E}_1$ and $\mathcal{E}_2$ in the circuit of **Fig. E26.24**, and find the potential difference of point b relative to point a.

Figure **E26.24**



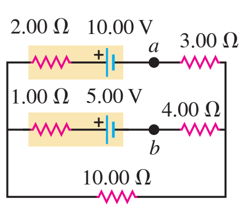
26.25 • In the circuit shown in **Fig. E26.25**, find (a) the current in the 3.00 Ω resistor; (b) the unknown emfs $\mathcal{E}_1$ and $\mathcal{E}_2$; (c) the resistance R. Note that three currents are given.

Figure **E26.25**



26.26 •• In the circuit shown in **Fig. E26.26**, find (a) the current in each branch and (b) the potential difference V_{ab} of point a relative to point b.

Figure **E26.26**

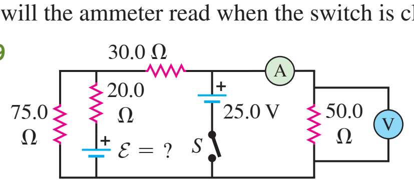


26.27 • The 10.00 V battery in Fig. E26.26 is removed from the circuit and reinserted with the opposite polarity, so that its positive terminal is now next to point a. The rest of the circuit is as shown in the figure. Find (a) the current in each branch and (b) the potential difference V_{ab} of point a relative to point b.

26.28 • The 5.00 V battery in Fig. E26.26 is removed from the circuit and replaced by a 15.00 V battery, with its negative terminal next to point b. The rest of the circuit is as shown in the figure. Find (a) the current in each branch and (b) the potential difference V_{ab} of point a relative to point b.

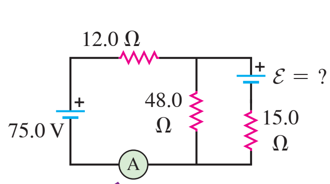
26.29 •• In the circuit shown in **Fig. E26.29** the batteries have negligible internal resistance and the meters are both idealized. With the switch S open, the voltmeter reads 15.0 V. (a) Find the emf $\mathcal{E}$ of the battery. (b) What will the ammeter read when the switch is closed?

Figure **E26.29**



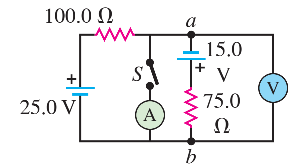
26.30 •• In the circuit shown in **Fig. E26.30** both batteries have insignificant internal resistance and the idealized ammeter reads 1.50 A in the direction shown. Find the emf $\mathcal{E}$ of the battery. Is the polarity shown correct?

Figure **E26.30**



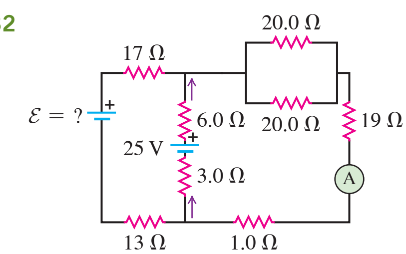
26.31 • In the circuit shown in **Fig. E26.31** all meters are idealized and the batteries have no appreciable internal resistance. (a) Find the reading of the voltmeter with the switch S open. Which point is at a higher potential: a or b? (b) With S closed, find the reading of the voltmeter and the ammeter. Which way (up or down) does the current flow through the switch?

Figure **E26.31**



26.32 •• In the circuit shown in **Fig. E26.32**, the 6.0 Ω resistor is consuming energy at a rate of 24 J/s when the current through it flows as shown. (a) Find the current through the ammeter A. (b) What are the polarity and emf $\mathcal{E}$ of the unknown battery, assuming it has negligible internal resistance?

Figure **E26.32**

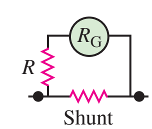


Section 26.3 Electrical Measuring Instruments

26.33 • The resistance of a galvanometer coil is 25.0 Ω, and the current required for full-scale deflection is 500 μA. (a) Show in a diagram how to convert the galvanometer to an ammeter reading 20.0 mA full scale, and compute the shunt resistance. (b) Show how to convert the galvanometer to a voltmeter reading 500 mV full scale, and compute the series resistance.

26.34 • The resistance of the coil of a pivoted-coil galvanometer is 9.36 Ω, and a current of 0.0224 A causes it to deflect full scale. We want to convert this galvanometer to an ammeter reading 20.0 A full scale. The only shunt available has a resistance of 0.0250 Ω. What resistance R must be connected in series with the coil (**Fig. E26.34**)?

Figure **E26.34**



26.35 • A circuit consists of a series combination of 6.00 kΩ and 5.00 kΩ resistors connected across a 50.0 V battery having negligible internal resistance. You want to measure the true potential difference (that is, the potential difference without the meter present) across the 5.00 kΩ resistor using a voltmeter having an internal resistance of 10.0 kΩ. (a) What potential difference does the voltmeter measure across the 5.00 kΩ resistor? (b) What is the *true* potential difference across this resistor when the meter is not present? (c) By what percentage is the voltmeter reading in error from the true potential difference?

Section 26.4 *R-C* Circuits

26.36 • To measure the capacitance C of a capacitor, you attach the capacitor to a battery and wait until it is fully charged. You then disconnect the capacitor from the battery and let it discharge through a resistor of resistance R. You measure the time $T_{1/2}$ that it takes the voltage across the resistor to decrease to half its initial value at the instant that the connection to the capacitor is first completed. You repeat this for several different resistors. You plot the data as $T_{1/2}$ versus R and find that they lie close to a straight line that has slope 5.00 μF. What is the capacitance C of the capacitor?

26.37 • A capacitor is charged to a potential of 12.0 V and is then connected to a voltmeter having a total internal resistance of 3.40 MΩ. After a time of 4.00 s the voltmeter reads 3.0 V. What are (a) the capacitance and (b) the time constant of the circuit?

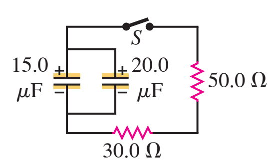
26.38 •• You connect a battery, resistor, and capacitor as in Fig. 26.20a, where $\mathcal{E} = 36.0$ V, $C = 5.00\ \mu$F, and $R = 120\ \Omega$. The switch S is closed at $t = 0$. (a) When the voltage across the capacitor is 8.00 V, what is the magnitude of the current in the circuit? (b) At what time t after the switch is closed is the voltage across the capacitor 8.00 V? (c) When the voltage across the capacitor is 8.00 V, at what rate is energy being stored in the capacitor?

26.39 • A 4.60 μF capacitor that is initially uncharged is connected in series with a 7.50 kΩ resistor and an emf source with $\mathcal{E} = 245$ V and negligible internal resistance. Just after the circuit is completed, what are (a) the voltage drop across the capacitor; (b) the voltage drop across the resistor; (c) the charge on the capacitor; (d) the current through the resistor? (e) A long time after the circuit is completed (after many time constants) what are the values of the quantities in parts (a)–(d)?

26.40 •• You connect a battery, resistor, and capacitor as in Fig. 26.20a, where $R = 12.0\ \Omega$ and $C = 5.00 \times 10^{-6}$ F. The switch S is closed at $t = 0$. When the current in the circuit has magnitude 3.00 A, the charge on the capacitor is 40.0×10^{-6} C. (a) What is the emf of the battery? (b) At what time t after the switch is closed is the charge on the capacitor equal to 40.0×10^{-6} C? (c) When the current has magnitude 3.00 A, at what rate is energy being (i) stored in the capacitor, (ii) supplied by the battery?

26.41 •• **CP** In the circuit shown in **Fig. E26.41** both capacitors are initially charged to 45.0 V. (a) How long after closing the switch S will the potential across each capacitor be reduced to 10.0 V, and (b) what will be the current at that time?

Figure **E26.41**



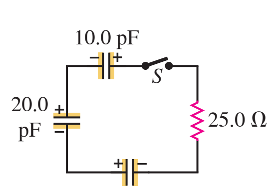
26.42 • A 12.4 μF capacitor is connected through a 0.895 MΩ resistor to a constant potential difference of 60.0 V. (a) Compute the charge on the capacitor at the following times after the connections are made: 0, 5.0 s, 10.0 s, 20.0 s, and 100.0 s. (b) Compute the charging currents at the same instants. (c) Graph the results of parts (a) and (b) for t between 0 and 20 s.

26.43 • An emf source with $\mathcal{E} = 120$ V, a resistor with $R = 80.0\ \Omega$, and a capacitor with $C = 4.00\ \mu$F are connected in series. As the capacitor charges, when the current in the resistor is 0.900 A, what is the magnitude of the charge on each plate of the capacitor?

26.44 • A network of two identical capacitors, each with capacitance C, is charged through a resistor R using a battery with emf $\mathcal{E}$. (a) What is the time constant, in terms of R and C, for the charging circuit if the two capacitors are in series? (b) In parallel? (c) For which capacitor network, series or parallel, does the voltage across the resistor become 1% of its initial value in a shorter time?

26.45 •• **CP** In the circuit shown in **Fig. E26.45** each capacitor initially has a charge of magnitude 3.50 nC on its plates. After the switch S is closed, what will be the current in the circuit at the instant that the capacitors have lost 80.0% of their initial stored energy?

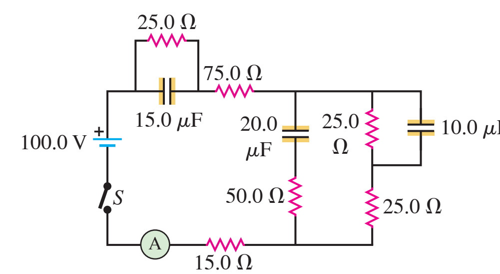
Figure **E26.45**



26.46 • A 1.50 μF capacitor is charging through a 12.0 Ω resistor using a 10.0 V battery. What will be the current when the capacitor has acquired $\frac{1}{4}$ of its maximum charge? Will it be $\frac{1}{4}$ of the maximum current?

26.47 • In the circuit in **Fig. E26.47** the capacitors are initially uncharged, the battery has no internal resistance, and the ammeter is idealized. Find the ammeter reading (a) just after the switch S is closed and (b) after S has been closed for a very long time.

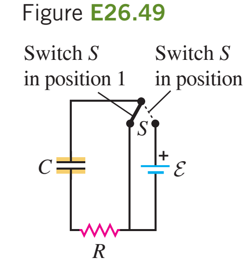
Figure **E26.47**



26.48 • A 12.0 μF capacitor is charged to a potential of 50.0 V and then discharged through a 225 Ω resistor. How long does it take the capacitor to lose (a) half of its charge and (b) half of its stored energy?

26.49 • In the circuit shown in **Fig. E26.49**, $C = 5.90\ \mu$F, $\mathcal{E} = 28.0$ V, and the emf has negligible resistance. Initially the capacitor is uncharged and the switch S is in position 1. The switch is then moved to position 2, so that the capacitor begins to charge. (a) What will be the charge on the capacitor a long time after S is moved to position 2? (b) After S has been in position 2 for 3.00 ms, the charge on the capacitor is measured to be 110 μC. What is the value of the resistance R? (c) How long after S is moved to position 2 will the charge on the capacitor be equal to 99.0% of the final value found in part (a)?

Figure **E26.49**



26.50 • A capacitor with $C = 6.00\ \mu$F is fully charged by connecting it to a battery that has emf 50.0 V. The capacitor is disconnected from the battery. A resistor of resistance $R = 185\ \Omega$ is connected across the capacitor, and the capacitor discharges through the resistor. (a) What is the charge q on the capacitor when the current in the resistor is 0.180 A? (b) If the connection to the resistor is completed at time $t = 0$, what is the value of t when the current has the value specified in part (a)?

Section 26.5 Power Distribution Systems

26.51 • A 1500 W electric heater is plugged into the outlet of a 120 V circuit that has a 20 A circuit breaker. You plug an electric hair dryer into the same outlet. The hair dryer has power settings of 600 W, 900 W, 1200 W, and 1500 W. You start with the hair dryer on the 600 W setting and increase the power setting until the circuit breaker trips. What power setting caused the breaker to trip?

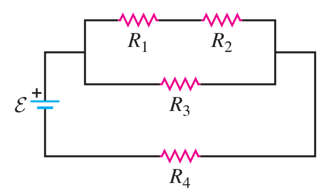
26.52 • The heating element of an electric dryer is rated at 4.1 kW when connected to a 240 V line. (a) What is the current in the heating element? Is 12 gauge wire large enough to supply this current? (b) What is the resistance of the dryer's heating element at its operating temperature? (c) At 11 cents per kWh, how much does it cost per hour to operate the dryer?

PROBLEMS

26.53 •• CALC A capacitor with capacitance C is connected in series to a resistor of resistance R and a battery with emf $\mathcal{E}$. The circuit is completed at time $t = 0$. (a) In terms of $\mathcal{E}$, R, and C, how much energy is stored in the capacitor when it is fully charged? (b) The power output of the battery is $P_{\mathcal{E}} = \mathcal{E}i$, with i given by Eq. (26.13). The electrical energy supplied in an infinitesimal time interval dt is $P_{\mathcal{E}}dt$. Integrate from $t = 0$ to $t \rightarrow \infty$ to find the total energy supplied by the battery. (c) The rate of consumption of electrical energy in the resistor is $P_R = i^2R$. In an infinitesimal time interval dt, the amount of electrical energy consumed by the resistor is $P_R dt$. Integrate from $t = 0$ to $t \rightarrow \infty$ to find the total energy consumed by the resistor. (d) What fraction of the total energy supplied by the battery is stored in the capacitor? What fraction is consumed in the resistor?

26.54 •• In **Fig. P26.54**, the battery has negligible internal resistance and $\mathcal{E} = 48.0$ V. $R_1 = R_2 = 4.00\ \Omega$ and $R_4 = 3.00\ \Omega$. What must the resistance R_3 be for the resistor network to dissipate electrical energy at a rate of 295 W?

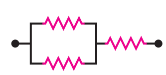
Figure P26.54



26.55 •• Consider the circuit shown in Fig. 26.6a. $\mathcal{E}_1 = 24.0$ V, $\mathcal{E}_2 = 36.0$ V, $r_1 = r_2 = 2.00\ \Omega$, and $R = 20.0\ \Omega$. (a) What is the current I_1 through $\mathcal{E}_1$? (b) What is the current I_2 through $\mathcal{E}_2$? (c) What is the current I_R through R?

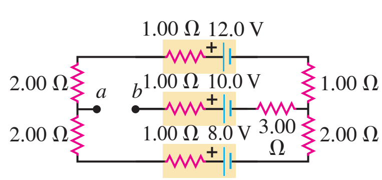
26.56 •• Each of the three resistors in **Fig. P26.56** has a resistance of 2.4 Ω and can dissipate a maximum of 48 W without becoming excessively heated. What is the maximum power the circuit can dissipate?

Figure P26.56



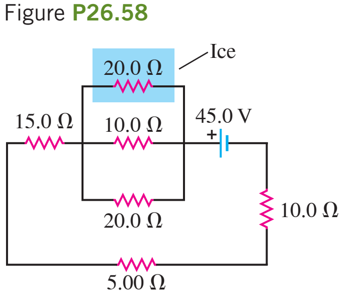
26.57 •• (a) Find the potential of point a with respect to point b in **Fig. P26.57**. (b) If points a and b are connected by a wire with negligible resistance, find the current in the 12.0 V battery.

Figure P26.57



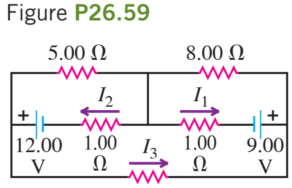
26.58 •• CP For the circuit shown in **Fig. P26.58** a 20.0 Ω resistor is embedded in a large block of ice at 0.00°C, and the battery has negligible internal resistance. At what rate (in g/s) is this circuit melting the ice? (The latent heat of fusion for ice is 3.34×10^5 J/kg.)

Figure P26.58



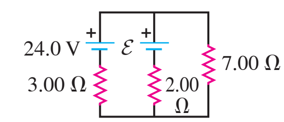
26.59 • Calculate the three currents I_1, I_2, and I_3 indicated in the circuit diagram shown in **Fig. P26.59**.

Figure P26.59



26.60 ••• What must the emf $\mathcal{E}$ in **Fig. P26.60** be in order for the current through the 7.00 Ω resistor to be 1.80 A? Each emf source has negligible internal resistance.

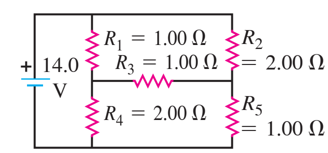
Figure P26.60



26.61 •• A standard phone charger provides 5.0 V of potential and supplies 1.0 A of current to the phone. (a) Estimate the diameter of each of the two wires inside your charge cable. (b) Assume the wire is copper, which has a resistivity of $1.72 \times 10^{-8}\ \Omega \cdot$m, and determine its resistance per meter. (c) If the phone draws $I = 1.0$ A, the potential supplied to the phone will be lower than the potential on the side of the charger by an amount IR, where R is the resistance of the wire. Determine the length of charging wire that would diminish the potential to 4.0 V at the phone.

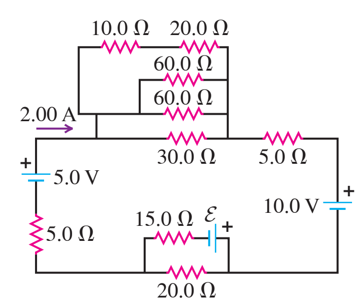
26.62 • (a) Find the current through the battery and each resistor in the circuit shown in **Fig. P26.62**. (b) What is the equivalent resistance of the resistor network?

Figure P26.62



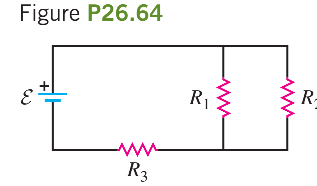
26.63 •• Consider the circuit shown in **Fig. P26.63**. (a) What must the emf $\mathcal{E}$ of the battery be in order for a current of 2.00 A to flow through the 5.00 V battery as shown? Is the polarity of the battery correct as shown? (b) How long does it take for 60.0 J of thermal energy to be produced in the 10.0 Ω resistor?

Figure P26.63



26.64 •• In the circuit shown in **Fig. P26.64**, $\mathcal{E} = 24.0$ V, $R_1 = 6.00\ \Omega$, $R_3 = 12.0\ \Omega$, and R_2 can vary between 3.00 Ω and 24.0 Ω. For what value of R_2 is the power dissipated by heating element R_1 the greatest? Calculate the magnitude of the greatest power.

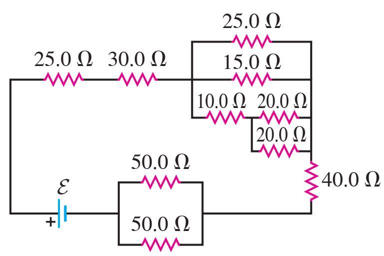
Figure P26.64



26.65 •• (a) Estimate energy usage and cost in your home by creating a table: In the first column list the household appliances that use electricity. In the second column indicate the voltage V used by each appliance. (Use 240 V for the hot-water heater, baseboard heat, and ovens; use 120 V for all other appliances.) In the third column list the current I drawn by each appliance. (This is readily obtained on the Internet.) In the fourth column indicate the number of daily hours T each appliance is used. In the fifth column compute the power of each appliance $P = IV$, and in the sixth column compute the daily energy use $U = PT$, in kilowatt-hours. At the bottom of the final column, sum the total kilowatt-hours used by these appliances. (b) Compute the monthly cost of electricity by multiplying the total daily kilowatt-hours by the cost (\$0.12/kW · h is typical). Then multiply the result by 30 days to obtain an estimate of your monthly cost of electricity.

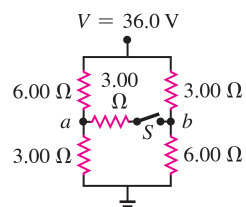
26.66 ••• In the circuit shown in **Fig. P26.66** all the resistors are rated at a maximum power of 2.00 W. What is the maximum emf $\mathcal{E}$ that the battery can have without burning up any of the resistors?

Figure **P26.66**



26.67 • **Figure P26.67** employs a convention often used in circuit diagrams. The battery (or other power supply) is not shown explicitly. It is understood that the point at the top, labeled "36.0 V," is connected to the positive terminal of a 36.0 V battery having negligible internal resistance, and that the *ground* symbol at the bottom is connected to the negative terminal of the battery. The circuit is completed through the battery, even though it is not shown. (a) What is the potential difference V_{ab}, the potential of point a relative to point b, when the switch S is open? (b) What is the current through S when it is closed? (c) What is the equivalent resistance when S is closed?

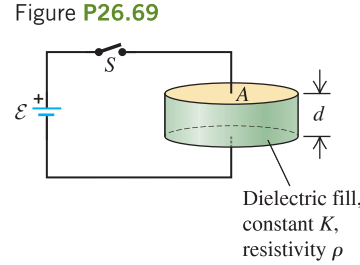
Figure **P26.67**



26.68 •• Three identical resistors are connected in series. When a certain potential difference is applied across the combination, the total power dissipated is 45.0 W. What power would be dissipated if the three resistors were connected in parallel across the same potential difference?

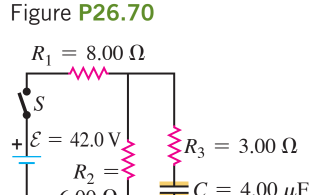
26.69 ••• CP Two capacitor plates with area A are separated by a distance d. The space between the plates is filled with dielectric material with dielectric constant K and resistivity ρ. This capacitor is attached to a battery that supplies a constant potential $\mathcal{E}$, as shown in **Fig. P26.69**, and is fully charged. At time $t = 0$ the switch S is opened. Owing to the nonzero value of ρ, this capacitor discharges by leaking. We can model this device as a capacitor and a resistor in parallel. (a) Analyze this circuit and determine the time constant, characterizing the discharge in terms of the parameters given above. Now assume the capacitor has plates with area $A = 1.00\ \text{cm}^2$ separated by $d = 30.0\ \mu\text{m}$ and is filled with a ceramic of dielectric constant $K = 8.70$ and resistivity $\rho = 3.10 \times 10^{12}\ \Omega \cdot \text{m}$. (b) If $\mathcal{E} = 5.00$ V, what is the charge on the capacitor at $t = 0$? (c) At what time will the capacitor have half of its original charge? (d) What is the magnitude of the leaking current at that point?

Figure **P26.69**



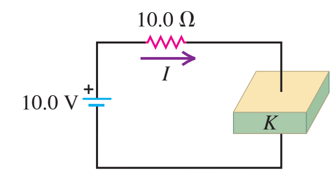
26.70 • The capacitor in **Fig. P26.70** is initially uncharged. The switch S is closed at $t = 0$. (a) Immediately after the switch is closed, what is the current through each resistor? (b) What is the final charge on the capacitor?

Figure **P26.70**



26.71 ••• CP CALC An R-C circuit is constructed as shown in **Fig. P26.71**. The capacitor has 2.00 cm × 2.00 cm square plates that are parallel and separated by a distance of 3.00 mm, and it is fully charged. At time $t = 0$, a 2.00 cm × 2.00 cm × 3.00 mm slab with dielectric constant $K = 12.0$ is slid between the plates as shown. (a) How much charge is on the upper plate before the slab is inserted? (b) What is the current I immediately after the slab is inserted? (c) How much energy is stored in the capacitor immediately after the slab is inserted? (d) By how much does the energy in the capacitor increase from that value after a long time has passed? (e) How much total energy is eventually supplied by the battery after the slab is inserted? (*Hint:* Determine the current as a function of time; then integrate the power delivered by the battery, $\int_0^\infty VI(t)\,dt$.) (f) How much energy is dissipated by the resistor?

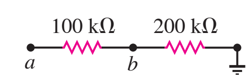
Figure **P26.71**



26.72 •• A 6.00 μF capacitor that is initially uncharged is connected in series with a 5.00 Ω resistor and an emf source with $\mathcal{E} = 50.0$ V and negligible internal resistance. At the instant when the resistor is dissipating electrical energy at a rate of 300 W, how much energy has been stored in the capacitor?

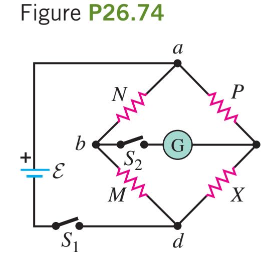
26.73 • Point a in **Fig. P26.73** is maintained at a constant potential of 400 V above ground. (See Problem 26.67.) (a) What is the reading of a voltmeter with the proper range and with resistance $5.00 \times 10^4\ \Omega$ when connected between point b and ground? (b) What is the reading of a voltmeter with resistance $5.00 \times 10^6\ \Omega$? (c) What is the reading of a voltmeter with infinite resistance?

Figure **P26.73**



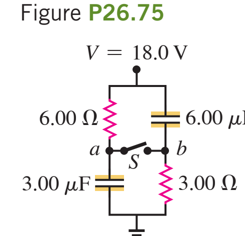
26.74 •• The Wheatstone Bridge. The circuit shown in **Fig. P26.74**, called a *Wheatstone bridge,* is used to determine the value of an unknown resistor X by comparison with three resistors M, N, and P whose resistances can be varied. For each setting, the resistance of each resistor is precisely known. With switches S_1 and S_2 closed, these resistors are varied until the current in the galvanometer G is zero; the bridge is then said to be *balanced.* (a) Show that under this condition the unknown resistance is given by $X = MP/N$. (This method permits very high precision in comparing resistors.) (b) If galvanometer G shows zero deflection when $M = 850.0\ \Omega$, $N = 15.00\ \Omega$, and $P = 33.48\ \Omega$, what is the unknown resistance X?

Figure **P26.74**



26.75 • (See Problem 26.67.) (a) What is the potential of point a with respect to point b in **Fig. P26.75** when the switch S is open? (b) Which point, a or b, is at the higher potential? (c) What is the final potential of point b with respect to ground when S is closed? (d) How much does the charge on each capacitor change when S is closed?

Figure **P26.75**

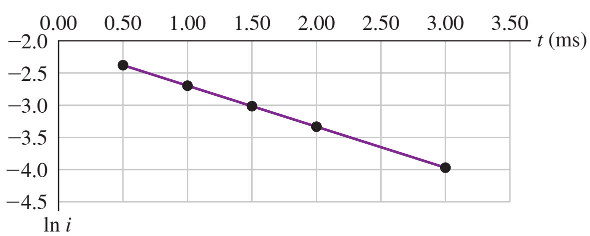


26.76 • A resistor with $R = 850\ \Omega$ is connected to the plates of a charged capacitor with capacitance $C = 4.62\ \mu$F. Just before the connection is made, the charge on the capacitor is 6.90 mC. (a) What is the energy initially stored in the capacitor? (b) What is the electrical power dissipated in the resistor just after the connection is made? (c) What is the electrical power dissipated in the resistor at the instant when the energy stored in the capacitor has decreased to half the value calculated in part (a)?

26.77 • A capacitor that is initially uncharged is connected in series with a resistor and an emf source with $\mathcal{E} = 110$ V and negligible internal resistance. Just after the circuit is completed, the current through the resistor is 6.5×10^{-5} A. The time constant for the circuit is 5.2 s. What are the resistance of the resistor and the capacitance of the capacitor?

26.78 •• **DATA** You set up the circuit shown in Fig. 26.22a, where $R = 196\ \Omega$. You close the switch at time $t = 0$ and measure the magnitude i of the current in the resistor R as a function of time t since the switch was closed. Your results are shown in **Fig. P26.78**, where you have chosen to plot $\ln i$ as a function of t. (a) Explain why your data points lie close to a straight line. (b) Use the graph in Fig. P26.78 to calculate the capacitance C and the initial charge Q_0 on the capacitor. (c) When $i = 0.0500$ A, what is the charge on the capacitor? (d) When $q = 0.500 \times 10^{-4}$ C, what is the current in the resistor?

Figure **P26.78**



26.79 •• **DATA** You set up the circuit shown in Fig. 26.20, where $C = 5.00 \times 10^{-6}$ F. At time $t = 0$, you close the switch and then measure the charge q on the capacitor as a function of the current i in the resistor. Your results are given in the table:

i (mA)	56.0	48.0	40.0	32.0	24.0
q (μC)	10.1	19.8	30.2	40.0	49.9

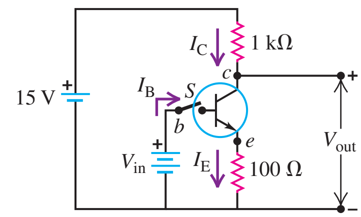
(a) Graph q as a function of i. Explain why the data points, when plotted this way, fall close to a straight line. Find the slope and y-intercept of the straight line that gives the best fit to the data. (b) Use your results from part (a) to calculate the resistance R of the resistor and the emf $\mathcal{E}$ of the battery. (c) At what time t after the switch is closed is the voltage across the capacitor equal to 10.0 V? (d) When the voltage across the capacitor is 4.00 V, what is the voltage across the resistor?

26.80 •• **DATA** The electronics supply company where you work has two different resistors, R_1 and R_2, in its inventory, and you must measure the values of their resistances. Unfortunately, stock is low, and all you have are R_1 and R_2 in parallel and in series—and you can't separate these two resistor combinations. You separately connect each resistor network to a battery with emf 48.0 V and negligible internal resistance and measure the power P supplied by the battery in both cases. For the series combination, $P = 48.0$ W; for the parallel combination, $P = 256$ W. You are told that $R_1 > R_2$. (a) Calculate R_1 and R_2. (b) For the series combination, which resistor consumes more power, or do they consume the same power? Explain. (c) For the parallel combination, which resistor consumes more power, or do they consume the same power?

CHALLENGE PROBLEMS

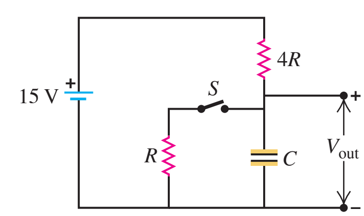
26.81 ••• The three-terminal device circled in **Fig. P26.81** is an NPN transistor. It works using three simple rules: I. The net current into the device is the same as the net current out of the device, so $I_C + I_B = I_E$. II. The potential at point e is always 0.60 V less than the potential at point b. Thus, $V_{in} - V_e = 0.60$ V. III. The current into point c is always a fixed multiple of the current into point b, so $I_C = \beta I_B$, where $\beta \gg 1$ is a parameter characteristic of the transistor. Use these rules to answer the questions. (a) What is the potential V_{out} in terms of V_{in}, in the limit $\beta \rightarrow \infty$? (*Hint:* Determine I_E in terms of V_{in}, and use this to determine I_C in terms of V_{in} and β. The potential difference across the 1.0 kΩ resistor can then be used to determine V_{out}. Take the limit $\beta \rightarrow \infty$ in this result.) (b) What value of V_{in} is needed so that $V_{out} = 7.5$ V? (c) If the input potential includes a small time-dependent "signal," so that $V_{in} = 15.0\text{ V} + v_{in}(t)$, then the output potential is $V_{out} = 7.5\text{ V} + Gv_{in}(t)$, where G is a "gain" factor. What is G for this circuit?

Figure **P26.81**



26.82 ••• In the circuit shown in **Fig. P26.82**, the switch S has been closed for a long time. (a) What is the output voltage V_{out}? Now the switch is opened and the output voltage increases. When the output voltage reaches 10.0 V, the switch closes and then the capacitor begins to discharge. When the output voltage reaches 5.0 V, the switch opens again. Thus this circuit undergoes periodic cycles of charging and discharging. (b) What is the time constant during a charging phase? (c) What is the time constant during a discharging phase? (d) What length of time passes between subsequent 10.0 V output voltages? (e) If $R = 10.0\text{ k}\Omega$ and $C = 10.0\ \mu$F, what is the frequency of operation? (Note: This circuit is a simplified example of a timer circuit, a crucial element in many electronics applications.)

Figure **P26.82**

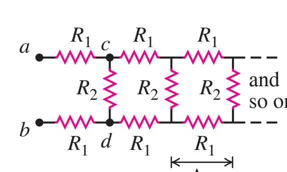


26.83 ••• **An Infinite Network.** As shown in **Fig. P26.83**, a network of resistors of resistances R_1 and R_2 extends to infinity toward the right. Prove that the total resistance R_T of the infinite network is equal to

$$R_T = R_1 + \sqrt{R_1^2 + 2R_1R_2}$$

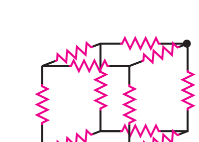
(*Hint:* Since the network is infinite, the resistance of the network to the right of points c and d is also equal to R_T.)

Figure **P26.83**



26.84 ••• Suppose a resistor R lies along each edge of a cube (12 resistors in all) with connections at the corners. Find the equivalent resistance between two diagonally opposite corners of the cube (points a and b in **Fig. P26.84**).

Figure **P26.84**



26.85 ••• **BIO Attenuator Chains and Axons.** The infinite network of resistors shown in Fig. P26.83 is known as an *attenuator chain,* since this chain of resistors causes the potential difference between the upper and lower wires to decrease, or attenuate, along the length of the chain. (a) Show that if the potential difference between the points a and b in Fig. 26.83 is V_{ab}, then the potential difference between points c and d is $V_{cd} = V_{ab}/(1 + \beta)$, where $\beta = 2R_1(R_T + R_2)/R_T R_2$ and R_T, the total resistance of the network, is given in Challenge Problem 26.83. (See the hint given in that problem.) (b) If the potential difference between terminals a and b at the left end of the infinite network is V_0, show that the potential difference between the upper and lower wires n segments from the left end is $V_n = V_0/(1 + \beta)^n$. If $R_1 = R_2$, how many segments are needed to decrease the potential difference V_n to less than 1.0% of V_0? (c) An infinite attenuator chain provides a model of the propagation of a voltage pulse along a nerve fiber, or axon. Each segment of the network in Fig. P26.83 represents a short segment of the axon of length Δx. The resistors R_1 represent the resistance of the fluid inside and outside the membrane wall of the axon. The resistance of the membrane to current flowing through the wall is represented by R_2. For an axon segment of length $\Delta x = 1.0\ \mu\text{m}$, $R_1 = 6.4 \times 10^3\ \Omega$ and $R_2 = 8.0 \times 10^8\ \Omega$ (the membrane wall is a good insulator). Calculate the total resistance R_T and β for an infinitely long axon. (This is a good approximation, since the length of an axon is much greater than its width; the largest axons in the human nervous system are longer than 1 m but only about 10^{-7} m in radius.) (d) By what fraction does the potential difference between the inside and outside of the axon decrease over a distance of 2.0 mm? (e) The attenuation of the potential difference calculated in part (d) shows that the axon cannot simply be a passive, current-carrying electrical cable; the potential difference must periodically be reinforced along the axon's length. This reinforcement mechanism is slow, so a signal propagates along the axon at only about 30 m/s. In situations where faster response is required, axons are covered with a segmented sheath of fatty myelin. The segments are about 2 mm long, separated by gaps called the *nodes of Ranvier*. The myelin increases the resistance of a 1.0-μm-long segment of the membrane to $R_2 = 3.3 \times 10^{12}\ \Omega$. For such a myelinated axon, by what fraction does the potential difference between the inside and outside of the axon decrease over the distance from one node of Ranvier to the next? This smaller attenuation means the propagation speed is increased.

MCAT-STYLE PASSAGE PROBLEMS

BIO Nerve Cells and *R-C* Circuits. The portion of a nerve cell that conducts signals is called an *axon*. Many of the electrical properties of axons are governed by ion channels, which are protein molecules that span the axon's cell membrane. When open, each ion channel has a pore that is filled with fluid of low resistivity and connects the interior of the cell electrically to the medium outside the cell. In contrast, the lipid-rich cell membrane in which ion channels reside has very high resistivity.

26.86 Assume that a typical open ion channel spanning an axon's membrane has a resistance of $1 \times 10^{11}\ \Omega$. We can model this ion channel, with its pore, as a 12-nm-long cylinder of radius 0.3 nm. What is the resistivity of the fluid in the pore? (a) $10\ \Omega \cdot \text{m}$; (b) $6\ \Omega \cdot \text{m}$; (c) $2\ \Omega \cdot \text{m}$; (d) $1\ \Omega \cdot \text{m}$.

26.87 In a simple model of an axon conducting a nerve signal, ions move across the cell membrane through open ion channels, which act as purely resistive elements. If a typical current density (current per unit cross-sectional area) in the cell membrane is $5\ \text{mA/cm}^2$ when the voltage across the membrane (the *action potential*) is 50 mV, what is the number density of open ion channels in the membrane? (a) $1/\text{cm}^2$; (b) $10/\text{cm}^2$; (c) $10/\text{mm}^2$; (d) $100/\mu\text{m}^2$.

26.88 Cell membranes across a wide variety of organisms have a capacitance per unit area of $1\ \mu\text{F/cm}^2$. For the electrical signal in a nerve to propagate down the axon, the charge on the membrane "capacitor" must change. What time constant is required when the ion channels are open? (a) 1 μs; (b) 10 μs; (c) 100 μs; (d) 1 ms.

ANSWERS

Chapter Opening Question ?

(ii) The potential difference V is the same across resistors connected in parallel. However, there is a different current I through each resistor if the resistances R are different: $I = V/R$.

Key Example VARIATION Problems

VP26.2.1 (a) 7.00 Ω (b) 0.571 Ω (c) 2.33 Ω (d) 0.857 Ω
VP26.2.2 (a) 8.18 A (b) 6.00 A (c) 2.18 A (d) 2.18 A
VP26.2.3 (a) 54.6 W (b) 20.6 W (c) 18.0 W (d) 16.0 W
VP26.2.4 (a) 9.84 W (b) 5.98 W (c) 2.08 W (d) 1.78 W
VP26.7.1 (a) 1 A (b) 7 V (c) 12 W for 12 V battery, 4 W for 4 V battery
VP26.7.2 (a) 7.75 V (b) 1.25 Ω (c) 0.926 Ω
VP26.7.3 (a) 33.4 W (b) 18.8 W (c) 2.09 W (d) 52.2 W (e) 25.0 W
VP26.7.4 (a) $(2\mathcal{E}_1 + \mathcal{E}_2)/3R$ (b) $-(\mathcal{E}_1 + 2\mathcal{E}_2)/3R$ (c) $(\mathcal{E}_1 - \mathcal{E}_2)/3R$
VP26.13.1 (a) 2.88×10^{-5} C (b) 9.00×10^{-7}A (c) 32.0 s (d) 0.430 (e) 0.570
VP26.13.2 (a) 11.0 s (b) 1.36×10^{-7} A
VP26.13.3 (a) 1.32 MΩ (b) 5.21×10^{-7} A (c) 1.04×10^{-7} A
VP26.13.4 (a) 6.12×10^{-5} C (b) 9.89×10^{-6} C

Bridging Problem

(a) 9.39 J (b) 2.02×10^4 W (c) 4.65×10^{-4} s (d) 7.43×10^3 W

? The needle of a magnetic compass points north. This alignment is due to (i) a magnetic force on the needle; (ii) a magnetic torque on the needle; (iii) the magnetic field that the needle itself produces; (iv) both (i) and (ii); (v) all of (i), (ii), and (iii).

27 Magnetic Field and Magnetic Forces

LEARNING OUTCOMES

In this chapter, you'll learn...

27.1 The properties of magnets, and how magnets interact with each other.
27.2 The nature of the force that a moving charged particle experiences in a magnetic field.
27.3 How magnetic field lines are different from electric field lines.
27.4 How to analyze the motion of a charged particle in a magnetic field.
27.5 Some practical applications of magnetic fields in chemistry and physics.
27.6 How to analyze magnetic forces on current-carrying conductors.
27.7 How current loops behave when placed in a magnetic field.
27.8 How direct-current motors work.
27.9 How magnetic forces give rise to the Hall effect.

You'll need to review...

1.10 Vector product of two vectors.
3.4, 5.4 Uniform circular motion.
10.1 Torque.
21.6, 21.7 Electric field lines and electric dipole moment.
22.2, 22.3 Electric flux and Gauss's law.
25.1 Electric current.
26.3 Galvanometers.

Everybody uses magnetic forces. They are at the heart of electric motors, microwave ovens, loudspeakers, computer printers, and disk drives. The most familiar examples of magnetism are permanent magnets, which attract unmagnetized iron objects and can also attract or repel other magnets. A compass needle aligning itself with the earth's magnetism is an example of this interaction. But the *fundamental* nature of magnetism is the interaction of moving electric charges. Unlike electric forces, which act on electric charges whether they are moving or not, magnetic forces act only on *moving* charges.

We saw in Chapter 21 that the electric force arises in two stages: (1) a charge produces an electric field in the space around it, and (2) a second charge responds to this field. Magnetic forces also arise in two stages. First, a *moving* charge or a collection of moving charges (that is, an electric current) produces a *magnetic* field. Next, a second current or moving charge responds to this magnetic field, and so experiences a magnetic force.

In this chapter we study the second stage in the magnetic interaction—that is, how moving charges and currents *respond* to magnetic fields. In particular, we'll see how to calculate magnetic forces and torques, and we'll discover why magnets can pick up iron objects like paper clips. In Chapter 28 we'll complete our picture of the magnetic interaction by examining how moving charges and currents *produce* magnetic fields.

27.1 MAGNETISM

Magnetic phenomena were first observed at least 2500 years ago in fragments of magnetized iron ore found near the ancient city of Magnesia (now Manisa, in western Turkey). These fragments were what are now called **permanent magnets;** you probably have several permanent magnets on your refrigerator door at home. Permanent magnets were found to exert forces on each other as well as on pieces of iron that were not magnetized. It was discovered that when an iron rod is brought in contact with a natural magnet, the rod also becomes magnetized. When such a rod is floated on water or suspended by a string from its center, it tends to line itself up in a north-south direction. The needle of an ordinary compass is just such a piece of magnetized iron.

Before the relationship of magnetic interactions to moving charges was understood, the interactions of permanent magnets and compass needles were described in terms of *magnetic poles*. If a bar-shaped permanent magnet, or *bar magnet,* is free to rotate, one end points north. This end is called a *north pole* or *N pole;* the other end is a *south pole* or *S pole*. Opposite poles attract each other, and like poles repel each other (**Fig. 27.1**). An object that contains iron but is not itself magnetized (that is, it shows no tendency to point north or south) is attracted by *either* pole of a permanent magnet (**Fig. 27.2**). This is the attraction that acts between a magnet and the unmagnetized steel door of a refrigerator. By analogy to electric interactions, we describe the interactions in Figs. 27.1 and 27.2 by saying that a bar magnet sets up a *magnetic field* in the space around it and a second object responds to that field. A compass needle tends to align with the magnetic field at the needle's position.

The earth itself is a magnet. Its north geographic pole is close to a magnetic *south* pole, which is why the north pole of a compass needle points north. The earth's magnetic axis is not quite parallel to its geographic axis (the axis of rotation), so a compass reading deviates somewhat from geographic north. This deviation, which varies with location, is called *magnetic declination* or *magnetic variation*. Also, the magnetic field is not horizontal at most points on the earth's surface; its angle up or down is called *magnetic inclination*. At the magnetic poles the magnetic field is vertical.

Figure 27.3 is a sketch of the earth's magnetic field. The lines, called *magnetic field lines,* show the direction that a compass would point at each location; they are discussed in detail in Section 27.3. The direction of the field at any point can be defined as the direction of the force that the field would exert on a magnetic north pole. In Section 27.2 we'll describe a more fundamental way to define the direction and magnitude of a magnetic field.

Magnetic Poles Versus Electric Charge

The concept of magnetic poles may appear similar to that of electric charge, and north and south poles may seem analogous to positive and negative charges. But the analogy can be misleading. While isolated positive and negative charges exist, there is *no* experimental evidence that one isolated magnetic pole exists; poles always appear in pairs. If a bar

Figure 27.3 A sketch of the earth's magnetic field. The field, which is caused by currents in the earth's molten core, changes with time; geologic evidence shows that it reverses direction entirely at irregular intervals of 10^4 to 10^6 years.

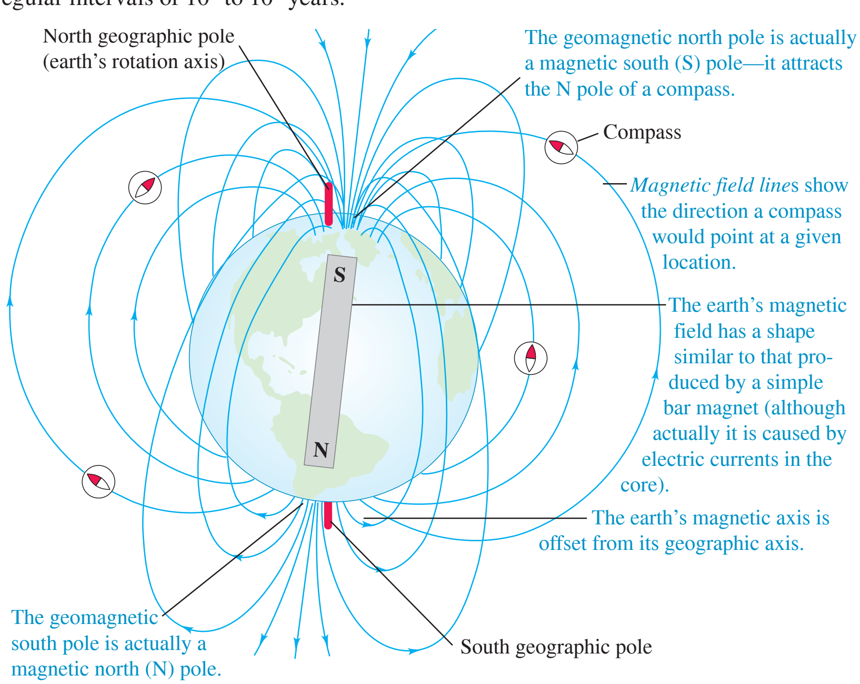


Figure 27.1 (a) Two bar magnets attract when opposite poles (N and S, or S and N) are next to each other. (b) The bar magnets repel when like poles (N and N, or S and S) are next to each other.

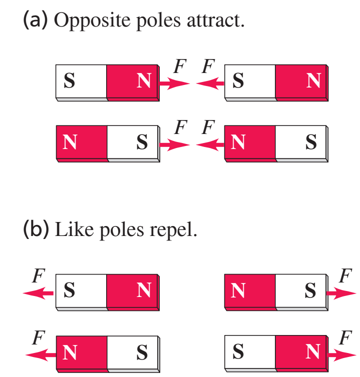


Figure 27.2 (a) Either pole of a bar magnet attracts an unmagnetized object that contains iron, such as a nail. (b) A real-life example of this effect.

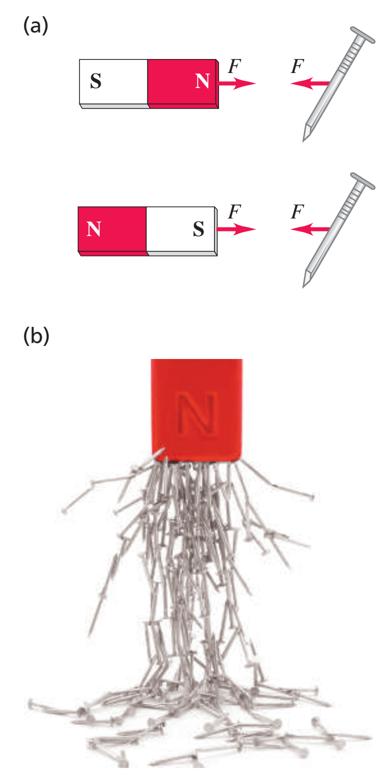


Figure **27.4** Breaking a bar magnet. Each piece has a north and south pole, even if the pieces are different sizes. (The smaller the piece, the weaker its magnetism.)

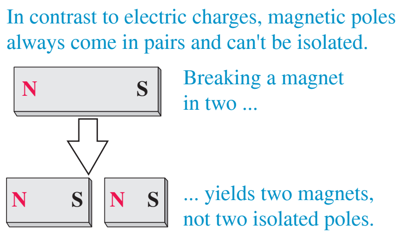
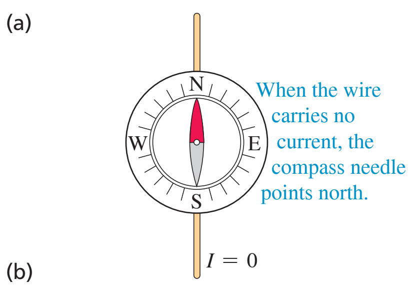
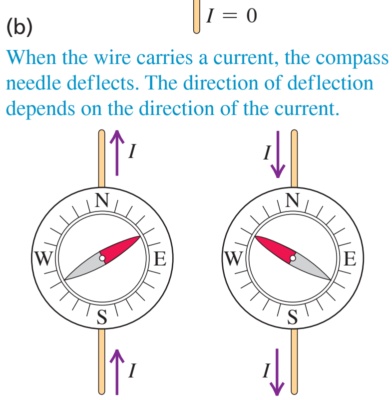


Figure **27.5** In Oersted's experiment, a compass is placed directly over a horizontal wire (here viewed from above).





BIO APPLICATION Spiny Lobsters and Magnetic Compasses Although the Caribbean spiny lobster (*Panulirus argus*) has a relatively simple nervous system, it is remarkably sensitive to magnetic fields. It has an internal magnetic "compass" that allows it to distinguish north, east, south, and west. This lobster can also sense small differences in the earth's magnetic field from one location to another and may use these differences to help it navigate.

magnet is broken in two, each broken end becomes a pole (**Fig. 27.4**). The existence of an isolated magnetic pole, or **magnetic monopole,** would have sweeping implications for theoretical physics. Extensive searches for magnetic monopoles have been carried out, but so far without success.

The first evidence of the relationship of magnetism to moving charges was discovered in 1820 by the Danish scientist Hans Christian Oersted. He found that a compass needle was deflected by a current-carrying wire (**Fig. 27.5**). Similar investigations were carried out in France by André Ampère. A few years later, Michael Faraday in England and Joseph Henry in the United States discovered that moving a magnet near a conducting loop can cause a current in the loop. We now know that the magnetic forces between two objects shown in Figs. 27.1 and 27.2 are fundamentally due to interactions between moving electrons in the atoms of the objects. (There are also *electric* interactions between the two objects, but these are far weaker than the magnetic interactions because the objects are electrically neutral.) Inside a magnetized object such as a permanent magnet, the motion of certain of the atomic electrons is *coordinated;* in an unmagnetized object these motions are not coordinated. (We'll describe these motions further in Section 27.7 and see how the interactions shown in Figs. 27.1 and 27.2 come about.)

Electric and magnetic interactions prove to be intimately connected. Over the next several chapters we'll develop the unifying principles of electromagnetism, culminating in the expression of these principles in *Maxwell's equations.* These equations represent the synthesis of electromagnetism, just as Newton's laws of motion are the synthesis of mechanics, and like Newton's laws they represent a towering achievement of the human intellect.

TEST YOUR UNDERSTANDING OF SECTION 27.1 Suppose you cut off the part of the compass needle shown in Fig. 27.5a that is painted gray. You discard this part, drill a hole in the remaining red part, and place the red part on the pivot at the center of the compass. Will the red part still swing when a current is applied as in Fig. 27.5b?

ANSWER

yes When a magnet is cut apart, each part has a north and south pole (see Fig. 27.4). Hence the small red part behaves much like the original, full-sized compass needle.

27.2 MAGNETIC FIELD

To introduce the concept of magnetic field properly, let's review our formulation of *electric* interactions in Chapter 21, where we introduced the concept of *electric* field. We represented electric interactions in two steps:

1. A distribution of electric charge creates an electric field $\vec{E}$ in the surrounding space.
2. The electric field exerts a force $\vec{F} = q\vec{E}$ on any other charge q that is present in the field.

We can describe magnetic interactions in a similar way:

1. A moving charge or a current creates a **magnetic field** in the surrounding space (in addition to its *electric* field).
2. The magnetic field exerts a force $\vec{F}$ on any other moving charge or current that is present in the field.

In this chapter we'll concentrate on the *second* aspect of the interaction: Given the presence of a magnetic field, what force does it exert on a moving charge or a current? In Chapter 28 we'll come back to the problem of how magnetic fields are *created* by moving charges and currents.

Like electric field, magnetic field is a *vector field*—that is, a vector quantity associated with each point in space. We'll use the symbol $\vec{B}$ for magnetic field. At any position the direction of $\vec{B}$ is defined as the direction in which the north pole of a compass needle tends to point. The arrows in Fig. 27.3 suggest the direction of the earth's magnetic field; for any magnet, $\vec{B}$ points out of its north pole and into its south pole.

Magnetic Forces on Moving Charges

There are four key characteristics of the magnetic force on a moving charge. First, its magnitude is proportional to the magnitude of the charge. If a 1 μC charge and a 2 μC charge move through a given magnetic field with the same velocity, experiments show that the force on the 2 μC charge is twice as great as the force on the 1 μC charge. Second, the magnitude of the force is also proportional to the magnitude, or "strength," of the field; if we double the magnitude of the field (for example, by using two identical bar magnets instead of one) without changing the charge or its velocity, the force doubles.

A third characteristic is that the magnetic force depends on the particle's velocity. This is quite different from the electric-field force, which is the same whether the charge is moving or not. A charged particle at rest experiences *no* magnetic force. And fourth, we find by experiment that the magnetic force $\vec{F}$ *does not* have the same direction as the magnetic field $\vec{B}$ but instead is always *perpendicular* to both $\vec{B}$ and the velocity $\vec{v}$. The magnitude F of the force is proportional to the component of $\vec{v}$ perpendicular to the field; when that component is zero (that is, when $\vec{v}$ and $\vec{B}$ are parallel or antiparallel), the force is zero.

Figure 27.6 shows these relationships. The direction of $\vec{F}$ is always perpendicular to the plane containing $\vec{v}$ and $\vec{B}$. Its magnitude is given by

$$F = |q|v_{\perp}B = |q|vB\sin\phi \tag{27.1}$$

where $|q|$ is the magnitude of the charge and ϕ is the angle measured from the direction of $\vec{v}$ to the direction of $\vec{B}$, as shown in the figure.

This description does not specify the direction of $\vec{F}$ completely; there are always two directions, opposite to each other, that are both perpendicular to the plane of $\vec{v}$ and $\vec{B}$. To complete the description, we use the same right-hand rule that we used to define the vector product in Section 1.10. (It would be a good idea to review that section before you go on.) Draw the vectors $\vec{v}$ and $\vec{B}$ with their tails together, as in **Fig. 27.7a**. Imagine turning $\vec{v}$ until it points in the direction of $\vec{B}$ (turning through the smaller of the two possible angles). Wrap the fingers of your right hand around the line perpendicular to the plane of $\vec{v}$ and $\vec{B}$ so that they curl around with the sense of rotation from $\vec{v}$ to $\vec{B}$. Your thumb then points in the direction of the force $\vec{F}$ on a *positive* charge.

This discussion shows that the force on a charge q moving with velocity $\vec{v}$ in a magnetic field $\vec{B}$ is given, both in magnitude and in direction, by

Magnetic force on a moving charged particle

$$\vec{F} = \overset{\text{Particle's charge}}{q}\underset{\text{Particle's velocity}}{\vec{v}} \times \underbrace{\vec{B}}_{\text{Magnetic field}} \tag{27.2}$$

Figure **27.6** The magnetic force $\vec{F}$ acting on a positive charge q moving with velocity $\vec{v}$ is perpendicular to both $\vec{v}$ and the magnetic field $\vec{B}$. For given values of speed v and magnetic field strength B, the force is greatest when $\vec{v}$ and $\vec{B}$ are perpendicular.

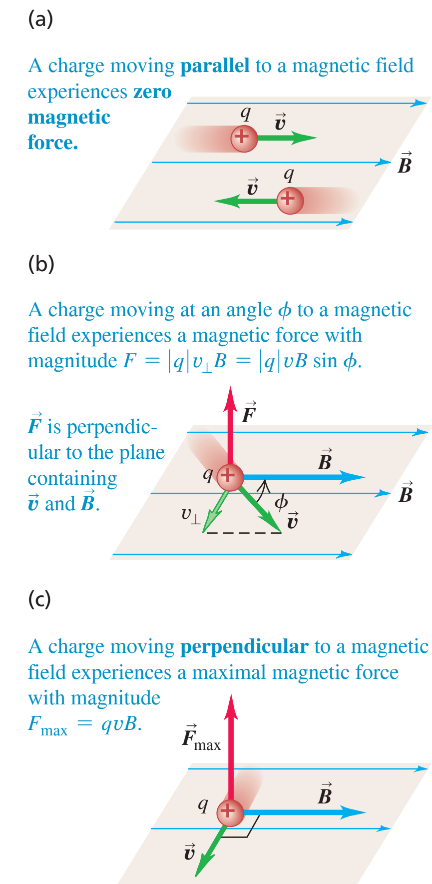


Figure **27.7** Finding the direction of the magnetic force on a moving charged particle.

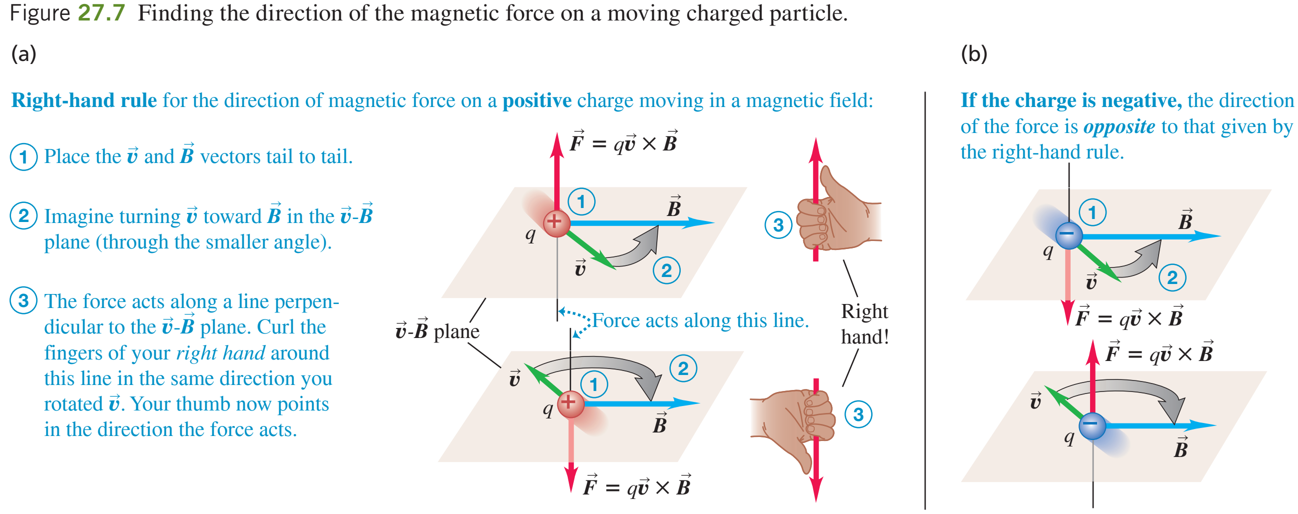
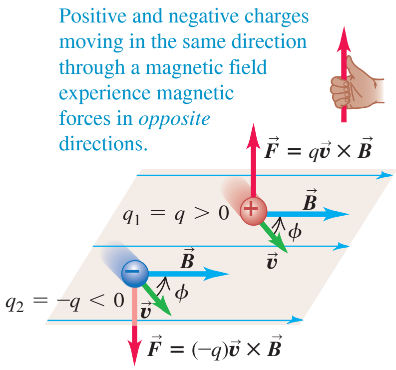


Figure 27.8 Reversing the sign of a moving charge reverses the direction of the magnetic force that acts on it.



This is the first of several vector products we'll encounter in our study of magnetic-field relationships. It's important to note that Eq. (27.2) was *not* deduced theoretically; it is an observation based on *experiment*.

Equation (27.2) is valid for both positive and negative charges. When q is negative, the direction of the force $\vec{F}$ is opposite to that of $\vec{v} \times \vec{B}$ (Fig. 27.7b). If two charges with equal magnitude and opposite sign move in the same $\vec{B}$ field with the same velocity (**Fig. 27.8**), the forces have equal magnitude and opposite direction. Figures 27.6, 27.7, and 27.8 show several examples of the relationships of the directions of $\vec{F}$, $\vec{v}$, and $\vec{B}$ for both positive and negative charges. Be sure you understand the relationships shown in these figures.

Equation (27.1) gives the magnitude of the magnetic force $\vec{F}$ in Eq. (27.2). Since ϕ is the angle between the directions of vectors $\vec{v}$ and $\vec{B}$, we may interpret $B \sin\phi$ as the component of $\vec{B}$ perpendicular to $\vec{v}$—that is, $B_\perp$. With this notation the force magnitude is

$$F = |q|vB_\perp \tag{27.3}$$

This form may be more convenient, especially in problems involving *currents* rather than individual particles. We'll discuss forces on currents later in this chapter.

From Eq. (27.1) the *units* of B must be the same as the units of F/qv. Therefore the SI unit of B is equivalent to $1\ \text{N} \cdot \text{s/C} \cdot \text{m}$, or, since one ampere is one coulomb per second $(1\ \text{A} = 1\ \text{C/s})$, $1\ \text{N/A} \cdot \text{m}$. This unit is called the **tesla** (abbreviated T), in honor of Nikola Tesla (1856–1943), the prominent Serbian-American scientist and inventor:

$$1\ \text{tesla} = 1\ \text{T} = 1\ \text{N/A} \cdot \text{m}$$

Another unit of B, the **gauss** $(1\ \text{G} = 10^{-4}\ \text{T})$, is also in common use.

The magnetic field of the earth is of the order of 10^{-4} T or 1 G. Magnetic fields of the order of 10 T occur in the interior of atoms and are important in the analysis of atomic spectra. The largest steady magnetic field that can be produced at present in the laboratory is about 45 T. Some pulsed-current electromagnets can produce fields of the order of 120 T for millisecond time intervals.

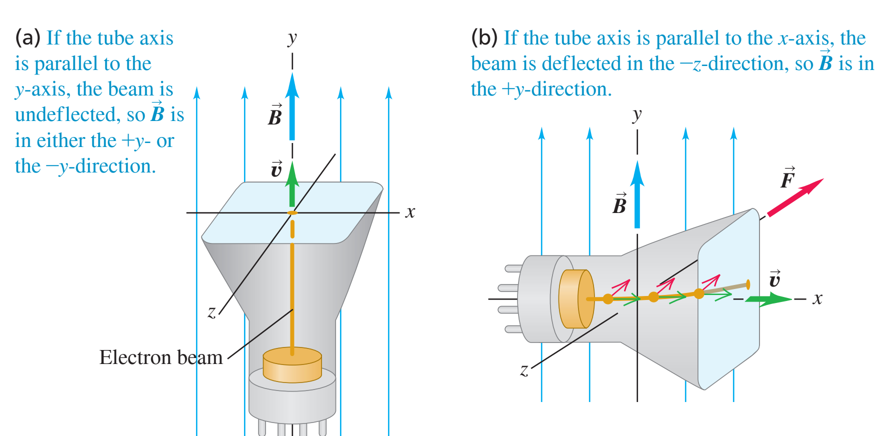
BIO APPLICATION Magnetic Fields of the Body All living cells are electrically active, and the feeble electric currents within your body produce weak but measurable magnetic fields. The fields produced by skeletal muscles have magnitudes less than 10^{-10} T, about one-millionth as strong as the earth's magnetic field. Your brain produces magnetic fields that are far weaker, only about 10^{-12} T.

Measuring Magnetic Fields with Test Charges

To explore an unknown magnetic field, we can measure the magnitude and direction of the force on a *moving* test charge and then use Eq. (27.2) to determine $\vec{B}$. The electron beam in a cathode-ray tube, such as that in an older television set (not a flat-screen TV), is a convenient device for this. The electron gun shoots out a narrow beam of electrons at a known speed. If there is no force to deflect the beam, it strikes the center of the screen.

If a magnetic field is present, in general the electron beam is deflected. But if the beam is parallel or antiparallel to the field, then $\phi = 0$ or π in Eq. (27.1) and $F = 0$; there is no force and hence no deflection. If we find that the electron beam is not deflected when its direction is parallel to a certain axis as in **Fig. 27.9a**, the $\vec{B}$ vector must point either up or down along that axis.

Figure 27.9 Determining the direction of a magnetic field by using a cathode-ray tube. Because electrons have a negative charge, the magnetic force $\vec{F} = q\vec{v} \times \vec{B}$ in part (b) points opposite to the direction given by the right-hand rule (see Fig. 27.7b).



If we then turn the tube 90° (Fig. 27.9b), $\phi = \pi/2$ in Eq. (27.1) and the magnetic force is maximum; the beam is deflected in a direction perpendicular to the plane of $\vec{B}$ and $\vec{v}$. The direction and magnitude of the deflection determine the direction and magnitude of $\vec{B}$. We can perform additional experiments in which the angle between $\vec{B}$ and $\vec{v}$ is between 0° and 90° to confirm Eq. (27.1). We note that the electron has a negative charge; the force in Fig. 27.9b is opposite in direction to the force on a positive charge.

When a charged particle moves through a region of space where *both* electric and magnetic fields are present, both fields exert forces on the particle. The total force $\vec{F}$ is the vector sum of the electric and magnetic forces:

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}) \tag{27.4}$$

CAUTION **No perpendicular field component, no magnetic force** When a charged particle moves in a magnetic field $\vec{B}$, the particle experiences a magnetic force *only* if there is a component of $\vec{B}$ that is perpendicular to the particle's velocity. If $\vec{B}$ and the velocity are in the same direction (as in Fig. 27.9a) or in opposite directions, there is no perpendicular component of $\vec{B}$ and hence no magnetic force.

PROBLEM-SOLVING STRATEGY 27.1 Magnetic Forces

IDENTIFY *the relevant concepts:* The equation $\vec{F} = q\vec{v} \times \vec{B}$ allows you to determine the magnetic force on a moving charged particle.

SET UP *the problem* using the following steps:

1. Draw the velocity $\vec{v}$ and magnetic field $\vec{B}$ with their tails together so that you can visualize the plane that contains them.
2. Determine the angle ϕ between $\vec{v}$ and $\vec{B}$.
3. Identify the target variables.

EXECUTE *the solution* as follows:

1. Use Eq. (27.2), $\vec{F} = q\vec{v} \times \vec{B}$, to express the magnetic force. Equation (27.1) gives the magnitude of the force, $F = qvB\sin\phi$.
2. Remember that $\vec{F}$ is perpendicular to the plane containing $\vec{v}$ and $\vec{B}$. The right-hand rule (see Fig. 27.7) gives the direction of $\vec{v} \times \vec{B}$. If q is negative, $\vec{F}$ is *opposite* to $\vec{v} \times \vec{B}$.

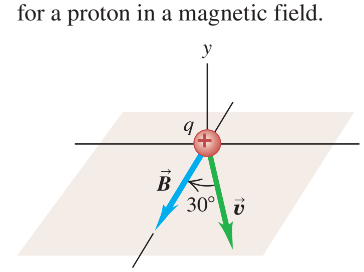
EVALUATE *your answer:* Whenever possible, solve the problem in two ways to confirm that the results agree. Do it directly from the geometric definition of the vector product. Then find the components of the vectors in some convenient coordinate system and calculate the vector product from the components. Verify that the results agree.

EXAMPLE 27.1 Magnetic force on a proton

WITH VARIATION PROBLEMS

A beam of protons ($q = 1.6 \times 10^{-19}$ C) moves at 3.0×10^5 m/s through a uniform 2.0 T magnetic field directed along the positive z-axis (**Fig. 27.10**). The velocity of each proton lies in the xz-plane and is directed at 30° to the $+z$-axis. Find the force on a proton.

Figure 27.10 Directions of $\vec{v}$ and $\vec{B}$ for a proton in a magnetic field.



IDENTIFY and SET UP This problem uses the expression $\vec{F} = q\vec{v} \times \vec{B}$ for the magnetic force $\vec{F}$ on a moving charged particle. The target variable is $\vec{F}$.

EXECUTE The charge is positive, so the force is in the same direction as the vector product $\vec{v} \times \vec{B}$. From the right-hand rule, this direction is along the negative y-axis. The magnitude of the force, from Eq. (27.1), is

$$\begin{aligned} F &= qvB\sin\phi \\ &= (1.6 \times 10^{-19}\ \text{C})(3.0 \times 10^5\ \text{m/s})(2.0\ \text{T})(\sin 30°) \\ &= 4.8 \times 10^{-14}\ \text{N} \end{aligned}$$

Continued

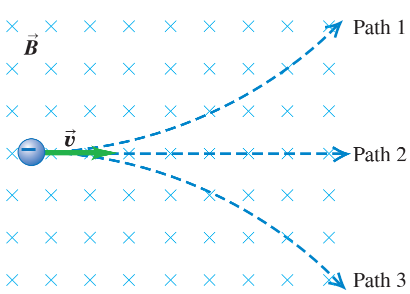
EVALUATE To check our result, we evaluate the force by using vector language and Eq. (27.2). We have

$$\begin{aligned}\vec{v} &= (3.0 \times 10^5\ \text{m/s})(\sin 30°)\hat{\imath} + (3.0 \times 10^5\ \text{m/s})(\cos 30°)\hat{k}\\ \vec{B} &= (2.0\ \text{T})\hat{k}\\ \vec{F} &= q\vec{v} \times \vec{B}\\ &= (1.6 \times 10^{-19}\ \text{C})(3.0 \times 10^5\ \text{m/s})(2.0\ \text{T})\\ &\quad \times (\sin 30°\hat{\imath} + \cos 30°\hat{k}) \times \hat{k}\\ &= (-4.8 \times 10^{-14}\ \text{N})\hat{\jmath}\end{aligned}$$

(Recall that $\hat{\imath} \times \hat{k} = -\hat{\jmath}$ and $\hat{k} \times \hat{k} = \mathbf{0}$.) We again find that the force is in the negative y-direction with magnitude 4.8×10^{-14} N.

If the beam consists of *electrons* rather than protons, the charge is negative ($q = -1.6 \times 10^{-19}$ C) and the direction of the force is reversed. The force is now directed along the *positive* y-axis, but the magnitude is the same as before, $F = 4.8 \times 10^{-14}$ N.

KEYCONCEPT A charged particle moving with velocity $\vec{v}$ through a magnetic field $\vec{B}$ experiences a magnetic force that depends on its charge, its speed, the magnitude of $\vec{B}$, and the angle between $\vec{v}$ and $\vec{B}$. The force is zero if the directions of $\vec{v}$ and $\vec{B}$ are the same or opposite; otherwise, the force is nonzero and in a direction perpendicular to both $\vec{v}$ and $\vec{B}$.



TEST YOUR UNDERSTANDING OF SECTION 27.2 The accompanying figure shows a uniform magnetic field $\vec{B}$ directed into the plane of the paper (shown by the blue ×'s). A particle with a negative charge moves in the plane. Which path—1, 2, or 3—does the particle follow?

ANSWER

path 3 Applying the right-hand rule to the vectors $\vec{v}$ (which points to the right) and $\vec{B}$ (which points into the plane of the figure) says that the force $\vec{F} = q\vec{v} \times \vec{B}$ on a *positive* charge would point *upward*. Since the charge is *negative*, the force points *downward* and the particle follows a trajectory that curves downward.

27.3 MAGNETIC FIELD LINES AND MAGNETIC FLUX

We can represent any magnetic field by **magnetic field lines,** just as we did for the earth's magnetic field in Fig. 27.3. The idea is the same as for the electric field lines we introduced in Section 21.6. We draw the lines so that the line through any point is tangent to the magnetic-field vector $\vec{B}$ at that point (**Fig. 27.11**). Just as with electric field lines, we draw only a few representative lines; otherwise, the lines would fill up all of space. Where adjacent field lines are close together, the field magnitude is large; where these field lines are far apart, the field magnitude is small. Also, because the direction of $\vec{B}$ at each point is unique, field lines never intersect.

CAUTION Magnetic field lines are not "lines of force" Unlike electric field lines, magnetic field lines *do not* point in the direction of the force on a charge (**Fig. 27.12**). Equation (27.2) shows that the force on a moving charged particle is always perpendicular to the magnetic field and hence to the magnetic field line that passes through the particle's position. The direction of the force depends on the particle's velocity and the sign of its charge, so just looking at magnetic field lines cannot tell you the direction of the force on an arbitrary moving charged particle. Magnetic field lines *do* have the direction that a compass needle would point at each location; this may help you visualize them.

Figure **27.11** Magnetic field lines of a permanent magnet. Note that the field lines pass through the interior of the magnet.

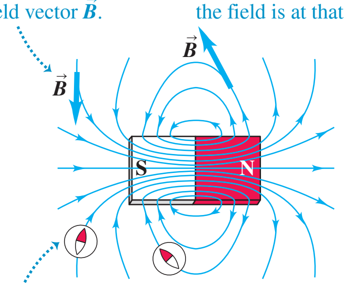


Figure **27.12** Magnetic field lines are *not* "lines of force."

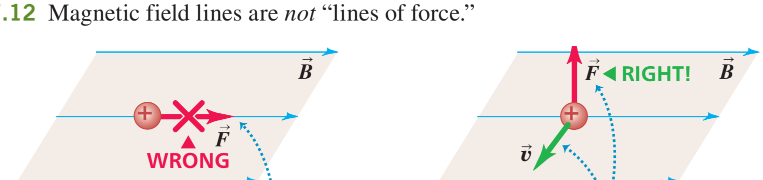
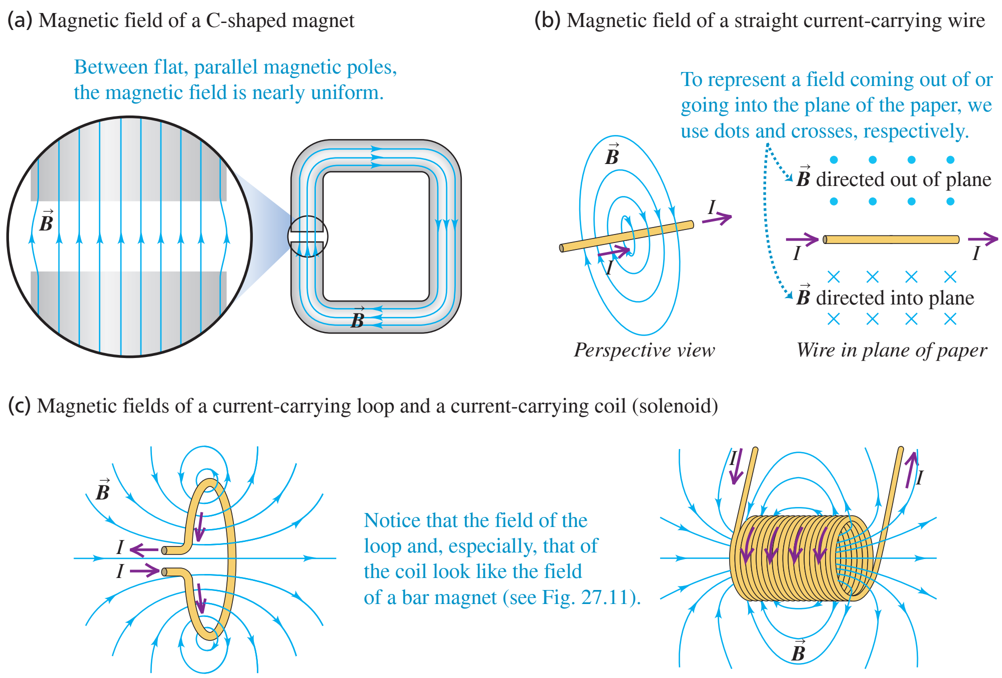


Figure **27.13** Magnetic field lines produced by some common sources of magnetic field.

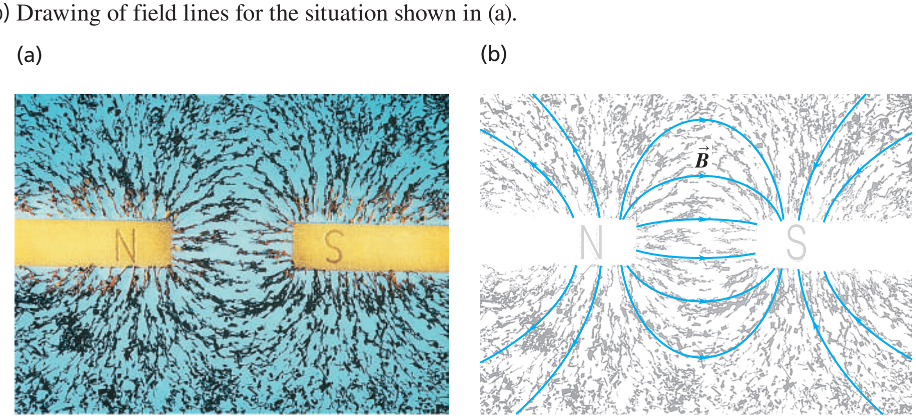


Figures 27.11 and 27.13 show magnetic field lines produced by several common sources of magnetic field. In the gap between the poles of the magnet shown in **Fig. 27.13a**, the field lines are approximately straight, parallel, and equally spaced, showing that the magnetic field in this region is approximately *uniform* (that is, constant in magnitude and direction).

Because magnetic-field patterns are three-dimensional, it's often necessary to draw magnetic field lines that point into or out of the plane of a drawing. To do this we use a dot ($\bullet$) to represent a vector directed out of the plane and a cross ($\times$) to represent a vector directed into the plane (Fig. 27.13b). To remember these, think of a dot as the head of an arrow coming directly toward you, and think of a cross as the feathers of an arrow flying directly away from you.

Iron filings, like compass needles, tend to align with magnetic field lines. Hence they provide an easy way to visualize field lines (**Fig. 27.14**).

Figure **27.14** (a) Like little compass needles, iron filings line up tangent to magnetic field lines. (b) Drawing of field lines for the situation shown in (a).



Magnetic Flux and Gauss's Law for Magnetism

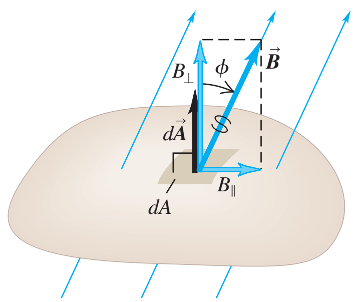


Figure 27.15 The magnetic flux through an area element dA is defined to be $d\Phi_B = B_\perp dA$.

We define the **magnetic flux** Φ_B through a surface just as we defined electric flux in connection with Gauss's law in Section 22.2. We can divide any surface into elements of area dA (**Fig. 27.15**). For each element we determine $B_\perp$, the component of $\vec{B}$ normal to the surface at the position of that element, as shown. From the figure, $B_\perp = B\cos\phi$, where ϕ is the angle between the direction of $\vec{B}$ and a line perpendicular to the surface. (Be careful not to confuse ϕ with Φ_B.) In general, this component varies from point to point on the surface. We define the magnetic flux $d\Phi_B$ through this area as

$$d\Phi_B = B_\perp dA = B\cos\phi\, dA = \vec{B}\cdot d\vec{A} \qquad (27.5)$$

The *total* magnetic flux through the surface is the sum of the contributions from the individual area elements:

$$\underbrace{\Phi_B}_{\text{Magnetic flux through a surface}} = \int \overbrace{B}^{\text{Magnitude of magnetic field } \vec{B}} \cos \underbrace{\phi}_{\text{Angle between } \vec{B} \text{ and normal to surface}} \underbrace{dA}_{\text{Element of surface area}} = \int \overbrace{B_\perp}^{\text{Component of } \vec{B} \text{ perpendicular to surface}} dA = \int \vec{B}\cdot \underbrace{d\vec{A}}_{\text{Vector element of surface area}} \qquad (27.6)$$

(Review the concepts of vector area and surface integral in Section 22.2.)

Magnetic flux is a *scalar* quantity. If $\vec{B}$ is uniform over a plane surface with total area A, then $B_\perp$ and ϕ are the same at all points on the surface, and

$$\Phi_B = B_\perp A = BA\cos\phi \qquad (27.7)$$

If $\vec{B}$ is also perpendicular to the surface (parallel to the area vector), then $\cos\phi = 1$ and Eq. (27.7) reduces to $\Phi_B = BA$. We'll use the concept of magnetic flux extensively during our study of electromagnetic induction in Chapter 29.

The SI unit of magnetic flux is equal to the unit of magnetic field (1 T) times the unit of area ($1\ \text{m}^2$). This unit is called the **weber** (1 Wb), in honor of the German physicist Wilhelm Weber (1804–1891):

$$1\ \text{Wb} = 1\ \text{T}\cdot\text{m}^2$$

Also, $1\ \text{T} = 1\ \text{N/A}\cdot\text{m}$, so

$$1\ \text{Wb} = 1\ \text{T}\cdot\text{m}^2 = 1\ \text{N}\cdot\text{m/A}$$

In Gauss's law the total *electric* flux through a closed surface is proportional to the total electric charge enclosed by the surface. For example, if the closed surface encloses an electric dipole, the total electric flux is zero because the total charge is zero. (You may want to review Section 22.3 on Gauss's law.) By analogy, if there were such a thing as a single magnetic charge (magnetic monopole), the total *magnetic* flux through a closed surface would be proportional to the total magnetic charge enclosed. But we have mentioned that no magnetic monopole has ever been observed, despite intensive searches. This leads us to **Gauss's law for magnetism:**

Gauss's law for magnetism:

$$\overbrace{\oint \vec{B}\cdot d\vec{A}}^{\text{The total magnetic flux through } any \text{ closed surface ...}} = \underbrace{0}_{\text{... equals zero.}} \qquad (27.8)$$

You can verify Gauss's law for magnetism by examining Figs. 27.11 and 27.13; if you draw a closed surface anywhere in any of the field maps shown in those figures, you'll see that every field line that enters the surface also exits from it; the net flux through the surface is zero. It also follows from Eq. (27.8) that magnetic field lines always form closed loops.

CAUTION Magnetic field lines have no ends Unlike electric field lines, which begin and end on electric charges, magnetic field lines *never* have endpoints; such a point would indicate the presence of a monopole. You might be tempted to draw magnetic field lines that begin at the north pole of a magnet and end at a south pole. But as Fig. 27.11 shows, a magnet's field lines continue through the interior of the magnet. Like all other magnetic field lines, they form closed loops.

For Gauss's law, which always deals with *closed* surfaces, the vector area element $d\vec{A}$ in Eq. (27.6) always points *out of* the surface. However, some applications of *magnetic* flux involve an *open* surface with a boundary line; there is then an ambiguity of sign in Eq. (27.6) because of the two possible choices of direction for $d\vec{A}$. In these cases we choose one of the two sides of the surface to be the "positive" side and use that choice consistently.

If the element of area dA in Eq. (27.5) is at right angles to the field lines, then $B_\perp = B$; calling the area $dA_\perp$, we have

$$B = \frac{d\Phi_B}{dA_\perp} \tag{27.9}$$

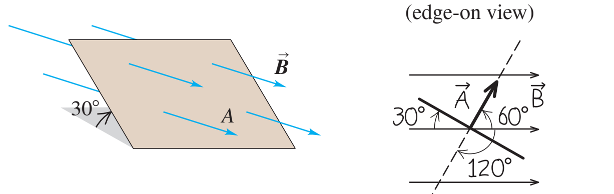
That is, the magnitude of the magnetic field is equal to *flux per unit area* across an area at right angles to the magnetic field. For this reason, magnetic field $\vec{B}$ is sometimes called **magnetic flux density.**

EXAMPLE 27.2 Magnetic flux calculations

Figure 27.16a is a perspective view of a flat surface with area 3.0 cm^2 in a uniform magnetic field $\vec{B}$. The magnetic flux through this surface is +0.90 mWb. Find the magnitude of the magnetic field and the direction of the area vector $\vec{A}$.

IDENTIFY and SET UP Our target variables are the field magnitude B and the direction of the area vector. Because $\vec{B}$ is uniform, B and ϕ are the same at all points on the surface. Hence we can use Eq. (27.7), $\Phi_B = BA\cos\phi$.

Figure 27.16 (a) A flat area A in a uniform magnetic field $\vec{B}$. (b) The area vector $\vec{A}$ makes a 60° angle with $\vec{B}$. (If we had chosen $\vec{A}$ to point in the opposite direction, ϕ would have been 120° and the magnetic flux Φ_B would have been negative.)



EXECUTE The area A is 3.0×10^{-4} m^2; the direction of $\vec{A}$ is perpendicular to the surface, so ϕ could be either 60° or 120°. But Φ_B, B, and A are all positive, so $\cos\phi$ must also be positive. This rules out 120°, so $\phi = 60°$ (Fig. 27.16b). Hence we find

$$B = \frac{\Phi_B}{A\cos\phi} = \frac{0.90 \times 10^{-3}\ \text{Wb}}{(3.0 \times 10^{-4}\ \text{m}^2)(\cos 60°)} = 6.0\ \text{T}$$

EVALUATE In many problems we are asked to calculate the flux of a given magnetic field through a given area. This example is somewhat different: It tests your understanding of the definition of magnetic flux.

KEYCONCEPT If a uniform magnetic field $\vec{B}$ is present over a surface, the magnetic flux through that surface equals the area of the surface multiplied by the component of $\vec{B}$ normal (perpendicular) to the surface. The net magnetic flux Φ_B through a closed surface is always zero.

TEST YOUR UNDERSTANDING OF SECTION 27.3 Imagine moving along the axis of the current-carrying loop in Fig. 27.13c, starting at a point well to the left of the loop and ending at a point well to the right of the loop. (a) How would the magnetic field strength vary as you moved along this path? (i) It would be the same at all points along the path; (ii) it would increase and then decrease; (iii) it would decrease and then increase. (b) Would the magnetic field direction vary as you moved along the path?

ANSWER

(a) **(ii)**, (b) **no** The magnitude of $\vec{B}$ would increase as you moved to the right, reaching a maximum as you passed through the plane of the loop. As you moved beyond the plane of the loop, the field magnitude would decrease. You can tell this from the spacing of the field lines: The closer the field lines, the stronger the field. The direction of the field would be to the right at all points along the path, since the path is along a field line and the direction of $\vec{B}$ at any point is tangent to the field line through that point.

27.4 MOTION OF CHARGED PARTICLES IN A MAGNETIC FIELD

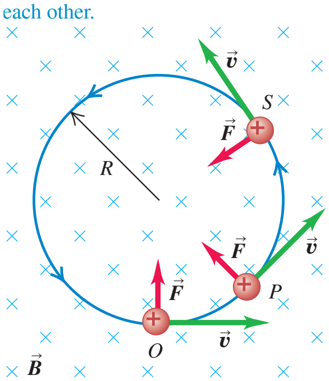
When a charged particle moves in a magnetic field, it is acted on by the magnetic force given by Eq. (27.2), and the motion is determined by Newton's laws. **Figure 27.17a** shows a simple example. A particle with positive charge q is at point O, moving with velocity $\vec{v}$ in a uniform magnetic field $\vec{B}$ directed into the plane of the figure. The vectors $\vec{v}$ and $\vec{B}$ are perpendicular, so the magnetic force $\vec{F} = q\vec{v} \times \vec{B}$ has magnitude $F = qvB$ and a direction as shown in the figure. The force is *always* perpendicular to $\vec{v}$, so it cannot change the *magnitude* of the velocity, only its direction. To put it differently, the magnetic force never has a component parallel to the particle's motion, so the magnetic force can never do *work* on the particle. This is true even if the magnetic field is not uniform.

Motion of a charged particle under the action of a magnetic field alone is always motion with constant speed.

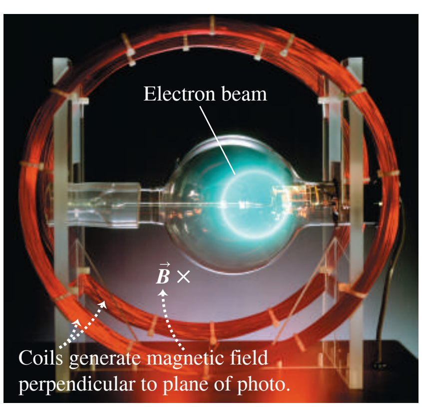
Figure 27.17 A charged particle moves in a plane perpendicular to a uniform magnetic field $\vec{B}$.

(a) The orbit of a charged particle in a uniform magnetic field

A charge moving at right angles to a uniform $\vec{B}$ field moves in a circle at constant speed because $\vec{F}$ and $\vec{v}$ are always perpendicular to each other.



(b) An electron beam (seen as a white arc) curving in a magnetic field



Using this principle, we see that in Fig. 27.17a the magnitudes of both $\vec{F}$ and $\vec{v}$ are constant. As the particle of mass m moves from O to P to S, the directions of force and velocity change but their magnitudes stay the same. So the particle is acted on by a constant-magnitude force that is always at right angles to the velocity of the particle. It follows from our discussion of circular motion in Sections 3.4 and 5.4 that the particle moves in a *circle* of radius R with constant speed v. The centripetal acceleration is v^2/R and only the magnetic force acts, so from Newton's second law,

$$F = |q|vB = m\frac{v^2}{R} \tag{27.10}$$

We solve Eq. (27.10) for R:

Particle's mass
Particle's speed
Radius of a circular orbit in a magnetic field
Magnetic-field magnitude
Particle's charge

$$R = \frac{mv}{|q|B} \tag{27.11}$$

If the charge q is negative, the particle moves *clockwise* around the orbit in Fig. 27.17a. Figure 27.17b shows a beam of negatively charged electrons following just such an orbit.

The angular speed ω of the particle can be found from Eq. (9.13), $v = R\omega$. Combining this with Eq. (27.11), we get

$$\omega = \frac{v}{R} = v\frac{|q|B}{mv} = \frac{|q|B}{m} \tag{27.12}$$

The number of revolutions per unit time is $f = \omega/2\pi$. This frequency f is independent of the radius R of the path. It is called the **cyclotron frequency;** in a particle accelerator called a *cyclotron,* particles moving in nearly circular paths are given a boost twice each revolution, increasing their energy and their orbital radii but not their angular speed or frequency. Similarly, one type of *magnetron,* a common source of microwave radiation for microwave ovens and radar systems, emits radiation with a frequency equal to the frequency of circular motion of electrons in a vacuum chamber between the poles of a magnet.

If the direction of the initial velocity is *not* perpendicular to the field, the velocity *component* parallel to the field is constant because there is no force parallel to the field. Then the particle moves in a helix (**Fig. 27.18**). The radius of the helix is given by Eq. (27.11), where v is now the component of velocity perpendicular to the $\vec{B}$ field.

Figure 27.18 The general case of a charged particle moving in a uniform magnetic field $\vec{B}$. The magnetic field does no work on the particle, so its speed and kinetic energy remain constant.

This particle's motion has components both parallel ($v_\parallel$) and perpendicular ($v_\perp$) to the magnetic field, so it moves in a helical path.

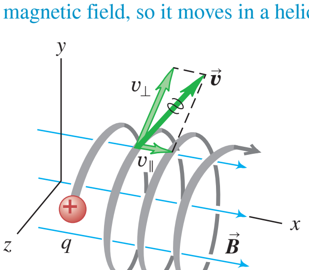


Figure 27.19 A magnetic bottle. Particles near either end of the region experience a magnetic force toward the center of the region. This is one way of containing an ionized gas that has a temperature of the order of 10^6 K, which would vaporize any material container.

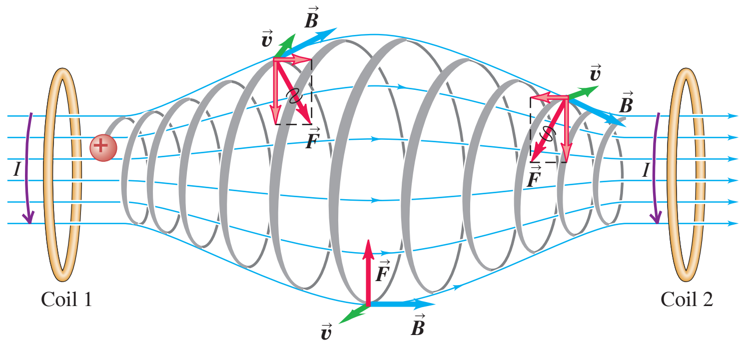
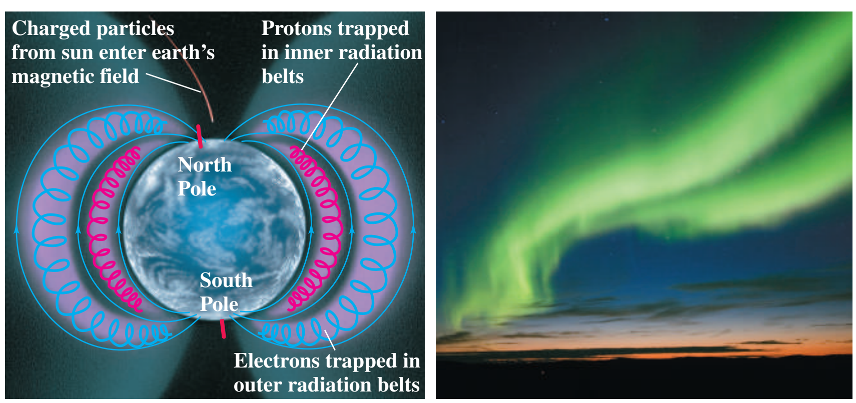


Figure 27.20 (a) The Van Allen radiation belts around the earth. Near the poles, charged particles from these belts can enter the atmosphere, producing the aurora borealis ("northern lights") and aurora australis ("southern lights"). (b) A photograph of the aurora borealis.

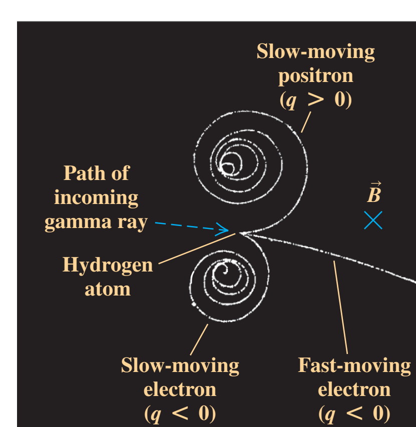
(a) (b)



Motion of a charged particle in a nonuniform magnetic field is more complex. **Figure 27.19** shows a field produced by two circular coils separated by some distance. Particles near either coil experience a magnetic force toward the center of the region; particles with appropriate speeds spiral repeatedly from one end of the region to the other and back. Because charged particles can be trapped in such a magnetic field, it is called a *magnetic bottle*. This technique is used to confine very hot plasmas with temperatures of the order of 10^6 K. In a similar way the earth's nonuniform magnetic field traps charged particles coming from the sun in doughnut-shaped regions around the earth, as shown in **Fig. 27.20**. These regions, called the *Van Allen radiation belts,* were discovered in 1958 from data obtained by instruments aboard the Explorer I satellite.

Magnetic forces on charged particles play an important role in studies of elementary particles. **Figure 27.21** shows a chamber filled with liquid hydrogen and with a magnetic field directed into the plane of the photograph. A high-energy gamma ray dislodges an electron from a hydrogen atom, sending it off at high speed and creating a visible track in the liquid hydrogen. The track shows the electron curving downward due to the magnetic force. The energy of the collision also produces another electron and a *positron* (a positively charged electron). Because of their opposite charges, the trajectories of the electron and the positron curve in opposite directions. As these particles plow through the liquid hydrogen, they collide with other charged particles, losing energy and speed. As a result, the radius of curvature decreases as suggested by Eq. (27.11). (The electron's speed is comparable to the speed of light, so Eq. (27.11) isn't directly applicable here.) Similar experiments allow physicists to determine the mass and charge of newly discovered particles.

Figure 27.21 This bubble chamber image shows the result of a high-energy gamma ray (which does not leave a track) that collides with an electron in a hydrogen atom. This electron flies off to the right at high speed. Some of the energy in the collision is transformed into a second electron and a positron (a positively charged electron). A magnetic field is directed into the plane of the image, which makes the positive and negative particles curve off in different directions.



PROBLEM-SOLVING STRATEGY 27.2 Motion in Magnetic Fields

IDENTIFY *the relevant concepts:* In analyzing the motion of a charged particle in electric and magnetic fields, you'll apply Newton's second law of motion, $\sum \vec{F} = m\vec{a}$, with the net force given by $\sum \vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$. Often other forces such as gravity can be ignored. Many of the problems are similar to the trajectory and circular-motion problems in Sections 3.3, 3.4, and 5.4; it would be a good idea to review those sections.

SET UP *the problem* using the following steps:

1. Determine the target variable(s).
2. Often the use of components is the most efficient approach. Choose a coordinate system and then express all vector quantities in terms of their components in this system.

EXECUTE *the solution* as follows:

1. If the particle moves perpendicular to a uniform magnetic field, the trajectory is a circle with a radius and angular speed given by Eqs. (27.11) and (27.12), respectively.
2. If your calculation involves a more complex trajectory, use $\sum \vec{F} = m\vec{a}$ in component form: $\sum F_x = ma_x$, and so forth. This approach is particularly useful when both electric and magnetic fields are present.

EVALUATE *your answer:* Check whether your results are reasonable.

EXAMPLE 27.3 Electron motion in a magnetron

WITH VARIATION PROBLEMS

A magnetron in a microwave oven emits electromagnetic waves with frequency $f = 2450$ MHz. What magnetic field strength is required for electrons to move in circular paths with this frequency?

IDENTIFY and SET UP The problem refers to circular motion as shown in Fig. 27.17a. We use Eq. (27.12) to solve for the field magnitude B.

EXECUTE The angular speed that corresponds to the frequency f is $\omega = 2\pi f = (2\pi)(2450 \times 10^6\ \text{s}^{-1}) = 1.54 \times 10^{10}\ \text{s}^{-1}$. Then from Eq. (27.12),

$$B = \frac{m\omega}{|q|} = \frac{(9.11 \times 10^{-31}\ \text{kg})(1.54 \times 10^{10}\ \text{s}^{-1})}{1.60 \times 10^{-19}\ \text{C}} = 0.0877\ \text{T}$$

EVALUATE This is a moderate field strength, easily produced with a permanent magnet. Incidentally, 2450 MHz electromagnetic waves are useful for heating and cooking food because they are strongly absorbed by water molecules.

KEYCONCEPT A charged particle that travels perpendicular to a uniform magnetic field $\vec{B}$ undergoes uniform circular motion. The frequency of the motion is independent of the speed of the particle: It depends on only the magnetic field strength and the particle's charge and mass.

EXAMPLE 27.4 Helical particle motion in a magnetic field

WITH VARIATION PROBLEMS

In a situation like that shown in Fig. 27.18, the charged particle is a proton ($q = 1.60 \times 10^{-19}$ C, $m = 1.67 \times 10^{-27}$ kg) and the uniform, 0.500 T magnetic field is directed along the x-axis. At $t = 0$ the proton has velocity components $v_x = 1.50 \times 10^5$ m/s, $v_y = 0$, and $v_z = 2.00 \times 10^5$ m/s. Only the magnetic force acts on the proton. (a) At $t = 0$, find the force on the proton and its acceleration. (b) Find the radius of the resulting helical path, the angular speed of the proton, and the *pitch* of the helix (the distance traveled along the helix axis per revolution).

IDENTIFY and SET UP The magnetic force is $\vec{F} = q\vec{v} \times \vec{B}$; Newton's second law gives the resulting acceleration. Because $\vec{F}$ is perpendicular to $\vec{v}$, the proton's speed does not change. Hence Eq. (27.11) gives the radius of the helical path if we replace v with the velocity component perpendicular to $\vec{B}$. Equation (27.12) gives the angular speed ω, which yields the time T for one revolution (the *period*). Given the velocity component parallel to the magnetic field, we can then determine the pitch.

EXECUTE (a) With $\vec{B} = B\hat{\imath}$ and $\vec{v} = v_x\hat{\imath} + v_z\hat{k}$, Eq. (27.2) yields

$$\begin{aligned}\vec{F} &= q\vec{v} \times \vec{B} = q(v_x\hat{\imath} + v_z\hat{k}) \times B\hat{\imath} = qv_zB\hat{\jmath}\\ &= (1.60 \times 10^{-19}\ \text{C})(2.00 \times 10^5\ \text{m/s})(0.500\ \text{T})\hat{\jmath}\\ &= (1.60 \times 10^{-14}\ \text{N})\hat{\jmath}\end{aligned}$$

(Recall: $\hat{\imath} \times \hat{\imath} = \mathbf{0}$ and $\hat{k} \times \hat{\imath} = \hat{\jmath}$.) The resulting acceleration is

$$\vec{a} = \frac{\vec{F}}{m} = \frac{1.60 \times 10^{-14}\ \text{N}}{1.67 \times 10^{-27}\ \text{kg}}\hat{\jmath} = (9.58 \times 10^{12}\ \text{m/s}^2)\hat{\jmath}$$

(b) Since $v_y = 0$, the component of velocity perpendicular to $\vec{B}$ is v_z; then from Eq. (27.11),

$$\begin{aligned}R &= \frac{mv_z}{|q|B} = \frac{(1.67 \times 10^{-27}\ \text{kg})(2.00 \times 10^5\ \text{m/s})}{(1.60 \times 10^{-19}\ \text{C})(0.500\ \text{T})}\\ &= 4.18 \times 10^{-3}\ \text{m} = 4.18\ \text{mm}\end{aligned}$$

From Eq. (27.12) the angular speed is

$$\omega = \frac{|q|B}{m} = \frac{(1.60 \times 10^{-19}\ \text{C})(0.500\ \text{T})}{1.67 \times 10^{-27}\ \text{kg}} = 4.79 \times 10^7\ \text{rad/s}$$

The period is $T = 2\pi/\omega = 2\pi/(4.79 \times 10^7\ \text{s}^{-1}) = 1.31 \times 10^{-7}$ s. The pitch is the distance traveled along the x-axis in this time, or

$$\begin{aligned}v_xT &= (1.50 \times 10^5\ \text{m/s})(1.31 \times 10^{-7}\ \text{s})\\ &= 0.0197\ \text{m} = 19.7\ \text{mm}\end{aligned}$$

EVALUATE Although the magnetic force has a tiny magnitude, it produces an immense acceleration because the proton mass is so small. Note that the pitch of the helix is almost five times greater than the radius R, so this helix is much more "stretched out" than that shown in Fig. 27.18.

KEYCONCEPT The most general path that a charged particle follows in a uniform magnetic field $\vec{B}$ is a helix: straight-line, constant-speed motion parallel to $\vec{B}$ combined with uniform circular motion perpendicular to $\vec{B}$.

TEST YOUR UNDERSTANDING OF SECTION 27.4 (a) If you double the speed of the charged particle in Fig. 27.17a while keeping the magnetic field the same (as well as the charge and the mass), how does this affect the radius of the trajectory? (i) The radius is unchanged; (ii) the radius is twice as large; (iii) the radius is four times as large; (iv) the radius is $\frac{1}{2}$ as large; (v) the radius is $\frac{1}{4}$ as large. (b) How does this affect the time required for one complete circular orbit? (i) The time is unchanged; (ii) the time is twice as long; (iii) the time is four times as long; (iv) the time is $\frac{1}{2}$ as long; (v) the time is $\frac{1}{4}$ as long.

ANSWER

(a) **(ii)**, (b) **(i)** The radius of the orbit as given by Eq. (27.11) is directly proportional to the speed, so doubling the particle speed causes the radius to double as well. The particle has twice as far to travel to complete one orbit but is traveling at double the speed, so the time for one orbit is unchanged. This result also follows from Eq. (27.12), which states that the angular speed ω is independent of the linear speed v. Hence the time per orbit, $T = 2\pi/\omega$, likewise does not depend on v.

27.5 APPLICATIONS OF MOTION OF CHARGED PARTICLES

This section describes several applications of the principles introduced in this chapter. Study them carefully, watching for applications of Problem-Solving Strategy 27.2 (Section 27.4).

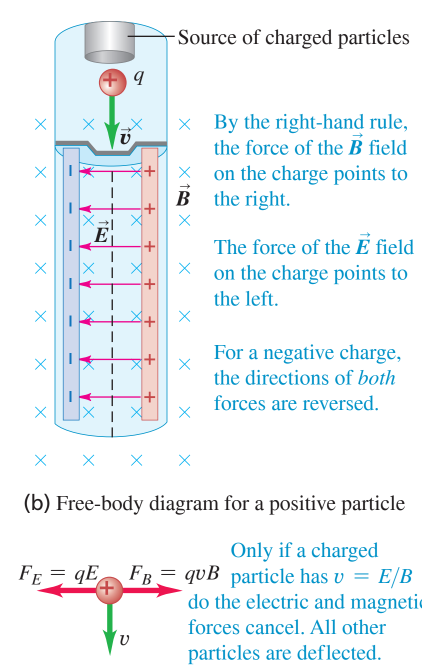
Velocity Selector

In a beam of charged particles produced by a heated cathode or a radioactive material, not all particles move with the same speed. Many applications, however, require a beam in which all the particle speeds are the same. Particles of a specific speed can be selected from the beam by using an arrangement of electric and magnetic fields called a *velocity selector*. In **Fig. 27.22a** a charged particle with mass m, charge q, and speed v enters a region of space where the electric and magnetic fields are perpendicular to the particle's velocity and to each other. The electric field $\vec{E}$ is to the left, and the magnetic field $\vec{B}$ is into the plane of the figure. If q is positive, the electric force is to the left, with magnitude qE, and the magnetic force is to the right, with magnitude qvB. For given field magnitudes E and B, for a particular value of v the electric and magnetic forces will be equal in magnitude; the total force is then zero, and the particle travels in a straight line with constant velocity. This will be the case if $qE = qvB$ (Fig. 27.22b), so the speed v for which there is no deflection is

$$v = \frac{E}{B} \tag{27.13}$$

Only particles with speeds equal to E/B can pass through without being deflected by the fields. By adjusting E and B appropriately, we can select particles having a particular speed for use in other experiments. Because q divides out in Eq. (27.13), a velocity selector for positively charged particles also works for electrons or other negatively charged particles.

Figure **27.22** (a) A velocity selector for charged particles uses perpendicular $\vec{E}$ and $\vec{B}$ fields. Only charged particles with $v = E/B$ move through undeflected. (b) The electric and magnetic forces on a positive charge. Both forces are reversed if the charge is negative.



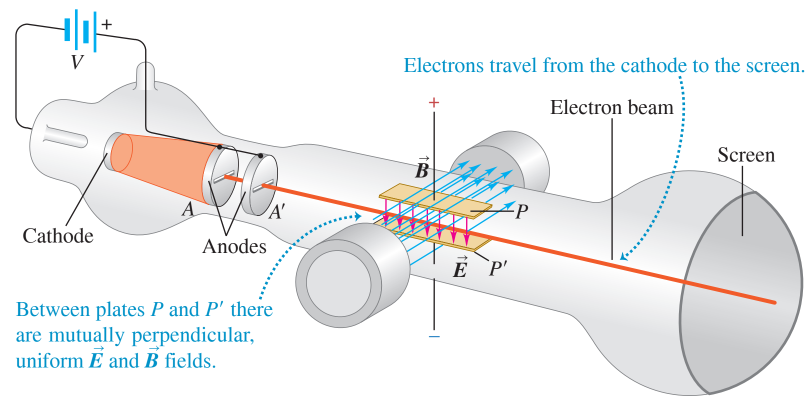
Thomson's e/m Experiment

In one of the landmark experiments in physics at the end of the 19th century, J. J. Thomson (1856–1940) used the idea just described to measure the ratio of charge to mass for the electron. For this experiment, carried out in 1897 at the Cavendish Laboratory in Cambridge, England, Thomson used the apparatus shown in **Fig. 27.23** (next page). In a highly evacuated glass container, electrons from the hot cathode are accelerated and formed into a beam by a potential difference V between the two anodes A and A'. The speed v of the electrons is determined by the accelerating potential V. The gained kinetic energy $\frac{1}{2}mv^2$ equals the lost electric potential energy eV, where e is the magnitude of the electron charge:

$$\frac{1}{2}mv^2 = eV \qquad \text{or} \qquad v = \sqrt{\frac{2eV}{m}} \tag{27.14}$$

The electrons pass between the plates P and P' and strike the screen at the end of the tube, which is coated with a material that fluoresces (glows) at the point of impact. The

Figure **27.23** Thomson's apparatus for measuring the ratio e/m for the electron.



electrons pass straight through the plates when Eq. (27.13) is satisfied; combining this with Eq. (27.14), we get

$$\frac{E}{B} = \sqrt{\frac{2eV}{m}} \qquad \text{so} \qquad \frac{e}{m} = \frac{E^2}{2VB^2} \tag{27.15}$$

All the quantities on the right side can be measured, so the ratio e/m of charge to mass can be determined. It is *not* possible to measure e or m separately by this method, only their ratio.

The most significant aspect of Thomson's e/m measurements was that he found a *single value* for this quantity. It did not depend on the cathode material, the residual gas in the tube, or anything else about the experiment. This independence showed that the particles in the beam, which we now call electrons, are a common constituent of all matter. Thus Thomson is credited with the first discovery of a subatomic particle, the electron.

The most precise value of e/m available as of this writing is

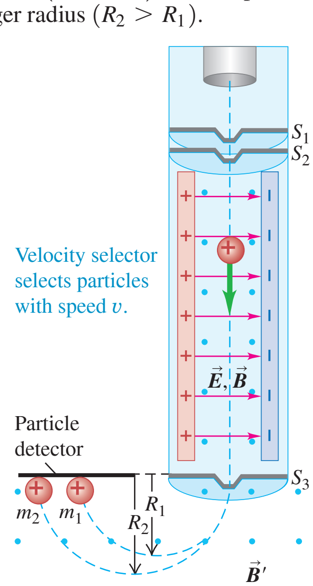
$$e/m = 1.758820024(11) \times 10^{11}\ \text{C/kg}$$

In this expression, (11) indicates the likely uncertainty in the last two digits, 24.

Fifteen years after Thomson's experiments, the American physicist Robert Millikan succeeded in measuring the charge of the electron precisely (see Challenge Problem 23.81). This value, together with the value of e/m, enables us to determine the *mass* of the electron. The most precise value available at present is

$$m = 9.10938356(11) \times 10^{-31}\ \text{kg}$$

Figure **27.24** Bainbridge's mass spectrometer utilizes a velocity selector to produce particles with uniform speed v. In the region of magnetic field B', particles with greater mass ($m_2 > m_1$) travel in paths with larger radius ($R_2 > R_1$).



Mass Spectrometers

Techniques similar to Thomson's e/m experiment can be used to measure masses of ions and thus measure atomic and molecular masses. In 1919, Francis Aston (1877–1945), a student of Thomson's, built the first of a family of instruments called **mass spectrometers.** A variation built by Bainbridge is shown in **Fig. 27.24**. Positive ions from a source pass through the slits S_1 and S_2, forming a narrow beam. Then the ions pass through a velocity selector with crossed $\vec{E}$ and $\vec{B}$ fields, as we have described, to block all ions except those with speeds v equal to E/B. Finally, the ions pass into a region with a magnetic field $\vec{B}'$ perpendicular to the figure, where they move in circular arcs with radius R determined by Eq. (27.11): $R = mv/qB'$. Ions with different masses strike the detector at different points, and the values of R can be measured. We assume that each ion has lost one electron, so the net charge of each ion is just $+e$. With everything known in this equation except m, we can compute the mass m of the ion.

One of the earliest results from this work was the discovery that neon has two species of atoms, with atomic masses 20 and 22 g/mol. We now call these species **isotopes** of the element. Later experiments have shown that many elements have several isotopes—atoms with identical chemical behaviors but different masses due to differing numbers of neutrons in their nuclei. This is just one of the many applications of mass spectrometers in chemistry and physics.

EXAMPLE 27.5 An e/m demonstration experiment

WITH VARIATION PROBLEMS

You set out to reproduce Thomson's e/m experiment with an accelerating potential of 150 V and a deflecting electric field of magnitude 6.0×10^6 N/C. (a) How fast do the electrons move? (b) What magnetic-field magnitude will yield zero beam deflection? (c) With this magnetic field, how will the electron beam behave if you increase the accelerating potential above 150 V?

IDENTIFY and SET UP This is the situation shown in Fig. 27.23. We use Eq. (27.14) to determine the electron speed and Eq. (27.13) to determine the required magnetic field B.

EXECUTE (a) From Eq. (27.14), the electron speed v is

$$v = \sqrt{2(e/m)V} = \sqrt{2(1.76 \times 10^{11}\ \text{C/kg})(150\ \text{V})}$$
$$= 7.27 \times 10^6\ \text{m/s} = 0.024c$$

(b) From Eq. (27.13), the required field strength is

$$B = \frac{E}{v} = \frac{6.0 \times 10^6\ \text{N/C}}{7.27 \times 10^6\ \text{m/s}} = 0.83\ \text{T}$$

(c) Increasing the accelerating potential V increases the electron speed v. In Fig. 27.23 this doesn't change the upward electric force eE, but it increases the downward magnetic force evB. Therefore the electron beam will turn *downward* and will hit the end of the tube below the undeflected position.

EVALUATE The required magnetic field is relatively large because the electrons are moving fairly rapidly (2.4% of the speed of light). If the maximum available magnetic field is less than 0.83 T, the electric field strength E would have to be reduced to maintain the desired ratio E/B in Eq. (27.15).

KEYCONCEPT A charged particle can experience both electric and magnetic forces if both an electric field $\vec{E}$ and a magnetic field $\vec{B}$ are present. These forces can cancel if $\vec{E}$, $\vec{B}$, and the particle's velocity are all mutually perpendicular and the particle travels at one particular speed [Eq. (27.14)].

EXAMPLE 27.6 Finding leaks in a vacuum system

WITH VARIATION PROBLEMS

There is almost no helium in ordinary air, so helium sprayed near a leak in a vacuum system will quickly show up in the output of a vacuum pump connected to such a system. You are designing a leak detector that uses a mass spectrometer to detect He^+ ions (charge $+e = +1.60 \times 10^{-19}$ C, mass 6.65×10^{-27} kg). Ions emerge from the velocity selector with a speed of 1.00×10^5 m/s. They are curved in a semicircular path by a magnetic field B' and are detected at a distance of 10.16 cm from the slit S_3 in Fig. 27.24. Calculate the magnitude of the magnetic field B'.

IDENTIFY and SET UP After it passes through the slit, the ion follows a circular path as described in Section 27.4 (see Fig. 27.17). We solve Eq. (27.11) for B'.

EXECUTE The distance given is the *diameter* of the semicircular path shown in Fig. 27.24, so the radius is $R = \frac{1}{2}(10.16 \times 10^{-2}\ \text{m})$. From Eq. (27.11), $R = mv/qB'$, we get

$$B' = \frac{mv}{qR} = \frac{(6.65 \times 10^{-27}\ \text{kg})(1.00 \times 10^5\ \text{m/s})}{(1.60 \times 10^{-19}\ \text{C})(5.08 \times 10^{-2}\ \text{m})} = 0.0818\ \text{T}$$

EVALUATE Helium leak detectors are widely used with high-vacuum systems. Our result shows that only a small magnetic field is required, so leak detectors can be relatively compact.

KEYCONCEPT The radius of the circular path that a charged particle follows in a uniform magnetic field depends on the particle's charge, mass, and speed. This dependence can be used to separate ions according to their mass.

TEST YOUR UNDERSTANDING OF SECTION 27.5 In Example 27.6 He^+ ions with charge $+e$ move at 1.00×10^5 m/s in a straight line through a velocity selector. Suppose the He^+ ions were replaced with He^{2+} ions, in which both electrons have been removed from the helium atom and the ion charge is $+2e$. At what speed must the He^{2+} ions travel through the same velocity selector in order to move in a straight line? (i) 4.00×10^5 m/s; (ii) 2.00×10^5 m/s; (iii) 1.00×10^5 m/s; (iv) 0.50×10^5 m/s; (v) 0.25×10^5 m/s.

ANSWER

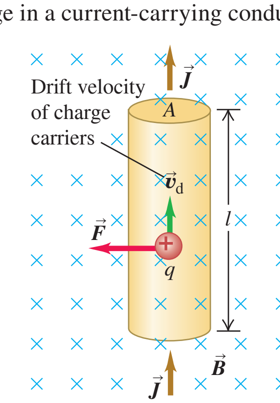
(iii) From Eq. (27.13), the speed $v = E/B$ at which particles travel straight through the velocity selector does not depend on the magnitude or sign of the charge or the mass of the particle. All that is required is that the particles (in this case, ions) have a nonzero charge.

27.6 MAGNETIC FORCE ON A CURRENT-CARRYING CONDUCTOR

What makes an electric motor work? Within the motor are conductors that carry currents (that is, whose charges are in motion), as well as magnets that exert forces on the moving charges. Hence there is a magnetic force on each current-carrying conductor, and these forces make the motor turn. The d'Arsonval galvanometer (Section 26.3) also uses magnetic forces on conductors.

We can compute the force on a current-carrying conductor starting with the magnetic force $\vec{F} = q\vec{v} \times \vec{B}$ on a single moving charge. **Figure 27.25** shows a straight segment of a conducting wire, with length l and cross-sectional area A; the current is from bottom to

Figure **27.25** Forces on a moving positive charge in a current-carrying conductor.



top. The wire is in a uniform magnetic field $\vec{B}$, perpendicular to the plane of the diagram and directed *into* the plane. Let's assume first that the moving charges are positive. Later we'll see what happens when they are negative.

The drift velocity $\vec{v}_d$ is upward, perpendicular to $\vec{B}$. The average force on each charge is $\vec{F} = q\vec{v}_d \times \vec{B}$, directed to the left as shown in the figure; since $\vec{v}_d$ and $\vec{B}$ are perpendicular, the magnitude of the force is $F = qv_dB$.

We can derive an expression for the *total* force on all the moving charges in a length l of conductor with cross-sectional area A by using the same language we used in Eqs. (25.2) and (25.3) of Section 25.1. The number of charges per unit volume, or charge concentration, is n; a segment of conductor with length l has volume Al and contains a number of charges equal to nAl. The total force $\vec{F}$ on *all* the moving charges in this segment has magnitude

$$F = (nAl)(qv_dB) = (nqv_dA)(lB) \tag{27.16}$$

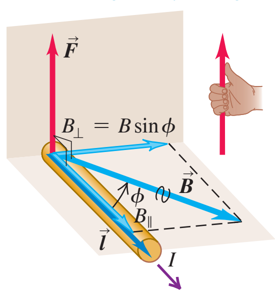
From Eq. (25.3) the current density is $J = nqv_d$. The product JA is the total current I, so we can rewrite Eq. (27.16) as

$$F = IlB \tag{27.17}$$

Figure **27.26** A straight wire segment of length $\vec{l}$ carries a current I in the direction of $\vec{l}$. The magnetic force on this segment is perpendicular to both $\vec{l}$ and the magnetic field $\vec{B}$.

Force $\vec{F}$ on a straight wire carrying a positive current and oriented at an angle ϕ to a magnetic field $\vec{B}$:

- Magnitude is $F = IlB_\perp = IlB\sin\phi$.
- Direction of $\vec{F}$ is given by the right-hand rule.



If the $\vec{B}$ field is not perpendicular to the wire but makes an angle ϕ with it, as in **Fig. 27.26**, we handle the situation the same way we did in Section 27.2 for a single charge. Only the component of $\vec{B}$ perpendicular to the wire (and to the drift velocities of the charges) exerts a force; this component is $B_\perp = B\sin\phi$. The magnetic force on the wire segment is then

$$F = IlB_\perp = IlB\sin\phi \tag{27.18}$$

The force is always perpendicular to both the conductor and the field, with the direction determined by the same right-hand rule we used for a moving positive charge (Fig. 27.26). Hence this force can be expressed as a vector product, like the force on a single moving charge. We represent the segment of wire with a vector $\vec{l}$ along the wire in the direction of the current; then the force $\vec{F}$ on this segment is

Magnetic force on a straight wire segment

$$\vec{F} = I\vec{l} \times \vec{B} \tag{27.19}$$

Current: I; Vector length of segment (points in current direction): $\vec{l}$; Magnetic field: $\vec{B}$

Figure **27.27** Magnetic field $\vec{B}$, length $\vec{l}$, and force $\vec{F}$ vectors for a straight wire carrying a current I.

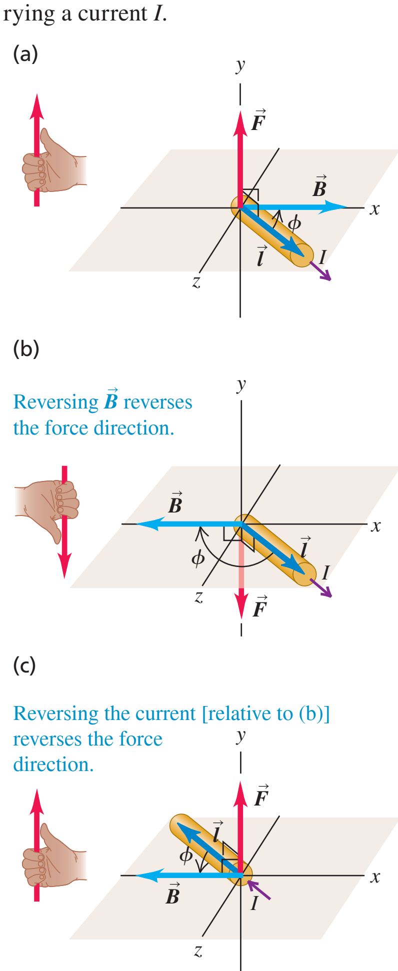


Figure 27.27 illustrates the directions of $\vec{B}$, $\vec{l}$, and $\vec{F}$ for several cases.

If the conductor is not straight, we can divide it into infinitesimal segments $d\vec{l}$. The force $d\vec{F}$ on each segment is

Magnetic force on an infinitesimal wire segment

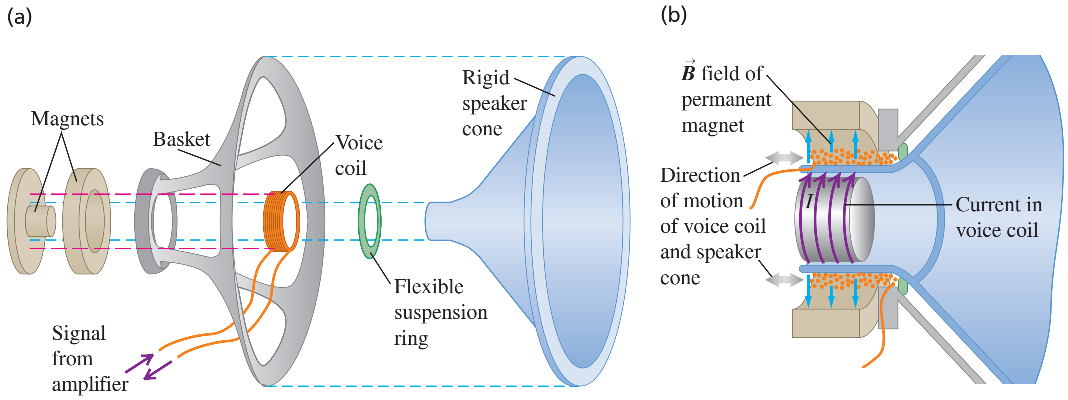
$$d\vec{F} = I\,d\vec{l} \times \vec{B} \tag{27.20}$$

Current: I; Vector length of segment (points in current direction): $d\vec{l}$; Magnetic field: $\vec{B}$

Then we can integrate this expression along the wire to find the total force on a conductor of any shape. The integral is a *line integral,* the same mathematical operation we have used to define work (Section 6.3) and electric potential (Section 23.2).

CAUTION **Current is not a vector** Recall from Section 25.1 that the current I is not a vector. The direction of current flow is described by $d\vec{l}$, not I. If the conductor is curved, I is the same at all points along its length, but $d\vec{l}$ changes direction—it is always tangent to the conductor.

Figure 27.28 (a) Components of a loudspeaker. (b) The permanent magnet creates a magnetic field that exerts forces on the current in the voice coil; for a current I in the direction shown, the force is to the right. If the electric current in the voice coil oscillates, the speaker cone attached to the voice coil oscillates at the same frequency.



Finally, what happens when the moving charges are negative, such as electrons in a metal? Then in Fig. 27.25 an upward current corresponds to a downward drift velocity. But because q is now negative, the direction of the force $\vec{F}$ is the same as before. Thus Eqs. (27.17) through (27.20) are valid for *both* positive and negative charges and even when *both* signs of charge are present at once. This happens in some semiconductor materials and in ionic solutions.

A common application of the magnetic forces on a current-carrying wire is found in loudspeakers (**Fig. 27.28**). The radial magnetic field created by the permanent magnet exerts a force on the voice coil that is proportional to the current in the coil; the direction of the force is either to the left or to the right, depending on the direction of the current. The signal from the amplifier causes the current to oscillate in direction and magnitude. The coil and the speaker cone to which it is attached respond by oscillating with an amplitude proportional to the amplitude of the current in the coil. Turning up the volume knob on the amplifier increases the current amplitude and hence the amplitudes of the cone's oscillation and of the sound wave produced by the moving cone.

EXAMPLE 27.7 Magnetic force on a straight conductor

WITH VARIATION PROBLEMS

A straight horizontal copper rod carries a current of 50.0 A from west to east in a region between the poles of a large electromagnet. In this region there is a horizontal magnetic field toward the northeast (that is, 45° north of east) with magnitude 1.20 T. (a) Find the magnitude and direction of the force on a 1.00 m section of rod. (b) While keeping the rod horizontal, how should it be oriented to maximize the magnitude of the force? What is the force magnitude in this case?

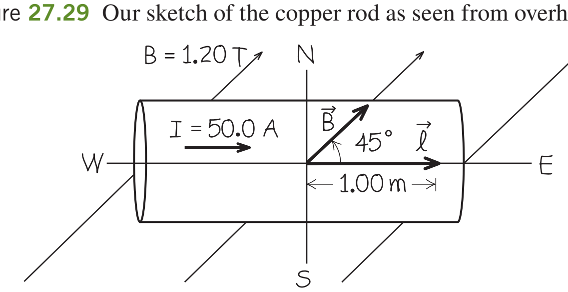


Figure 27.29 Our sketch of the copper rod as seen from overhead.

IDENTIFY and SET UP **Figure 27.29** shows the situation. This is a straight wire segment in a uniform magnetic field, as in Fig. 27.26. Our target variables are the force $\vec{F}$ on the segment and the angle ϕ for which the force magnitude F is greatest. We find the magnitude of the magnetic force from Eq. (27.18) and the direction from the right-hand rule.

EXECUTE (a) The angle ϕ between the directions of current and field is 45°. From Eq. (27.18) we obtain

$$F = IlB\sin\phi = (50.0\text{ A})(1.00\text{ m})(1.20\text{ T})(\sin 45°) = 42.4\text{ N}$$

The direction of the force is perpendicular to the plane of the current and the field, both of which lie in the horizontal plane. Thus the force must be vertical; the right-hand rule shows that it is vertically *upward* (out of the plane of the figure).

(b) From $F = IlB\sin\phi$, F is maximum for $\phi = 90°$, so that $\vec{l}$ and $\vec{B}$ are perpendicular. To keep $\vec{F} = I\vec{l} \times \vec{B}$ upward, we rotate the rod *clockwise* by 45° from its orientation in Fig. 27.29, so that the current runs toward the southeast. Then $F = IlB = (50.0\text{ A})(1.00\text{ m})(1.20\text{ T}) = 60.0\text{ N}$.

EVALUATE We check the result in part (a) by using Eq. (27.19) to calculate the force vector. If we use a coordinate system with the

Continued

x-axis pointing east, the y-axis north, and the z-axis upward, we have $\vec{l} = (1.00\ \text{m})\hat{\imath}$, $\vec{B} = (1.20\ \text{T})[(\cos 45°)\hat{\imath} + (\sin 45°)\hat{\jmath}]$, and

$$\begin{aligned}\vec{F} &= I\vec{l} \times \vec{B}\\ &= (50.0\ \text{A})(1.00\ \text{m})\hat{\imath} \times (1.20\ \text{T})[(\cos 45°)\hat{\imath} + (\sin 45°)\hat{\jmath}]\\ &= (42.4\ \text{N})\hat{k}\end{aligned}$$

Note that the maximum upward force of 60.0 N can hold the conductor in midair against the force of gravity—that is, *magnetically levitate* the conductor—if its weight is 60.0 N and its mass is $m = w/g = (60.0\ \text{N})/(9.8\ \text{m/s}^2) = 6.12$ kg. Magnetic levitation is used in some high-speed trains to suspend the train over the tracks. Eliminating rolling friction in this way allows the train to achieve speeds of over 400 km/h.

KEYCONCEPT A straight, current-carrying conductor in a magnetic field $\vec{B}$ experiences a magnetic force that depends on its current, its length, the magnitude of $\vec{B}$, and the angle between the current direction and $\vec{B}$. The force is zero if the directions of the current and $\vec{B}$ are the same or opposite; otherwise, the force is nonzero and in a direction perpendicular to both the current and $\vec{B}$.

EXAMPLE 27.8 Magnetic force on a curved conductor

In **Fig. 27.30** the magnetic field $\vec{B}$ is uniform and perpendicular to the plane of the figure, pointing out of the page. The conductor, carrying current I to the left, has three segments: (1) a straight segment with length L perpendicular to the plane of the figure, (2) a semicircle with radius R, and (3) another straight segment with length L parallel to the x-axis. Find the total magnetic force on this conductor.

IDENTIFY and SET UP The magnetic field $\vec{B} = B\hat{k}$ is uniform, so we find the forces $\vec{F}_1$ and $\vec{F}_3$ on the straight segments (1) and (3) from Eq. (27.19). We divide the curved segment (2) into infinitesimal straight segments and find the corresponding force $d\vec{F}_2$ on each straight segment from Eq. (27.20). We then integrate to find $\vec{F}_2$. The total magnetic force on the conductor is then $\vec{F} = \vec{F}_1 + \vec{F}_2 + \vec{F}_3$.

EXECUTE For segment (1), $\vec{L} = -L\hat{k}$. Hence from Eq. (27.19), $\vec{F}_1 = I\vec{L} \times \vec{B} = \mathbf{0}$. For segment (3), $\vec{L} = -L\hat{\imath}$, so $\vec{F}_3 = I\vec{L} \times \vec{B} = I(-L\hat{\imath}) \times (B\hat{k}) = ILB\hat{\jmath}$.

For the curved segment (2), Fig. 27.30 shows a segment $d\vec{l}$ with length $dl = R\,d\theta$, at angle θ. The right-hand rule shows that the direction of $d\vec{l} \times \vec{B}$ is radially outward from the center; make sure you can verify this. Because $d\vec{l}$ and $\vec{B}$ are perpendicular, the magnitude dF_2 of the force on the segment $d\vec{l}$ is $dF_2 = I\,dl\,B = I(R\,d\theta)B$. The components of the force on this segment are

$$dF_{2x} = IR\,d\theta\,B\cos\theta \qquad dF_{2y} = IR\,d\theta\,B\sin\theta$$

To find the components of the total force, we integrate these expressions with respect to θ from $\theta = 0$ to $\theta = \pi$ to take in the whole semicircle. The results are

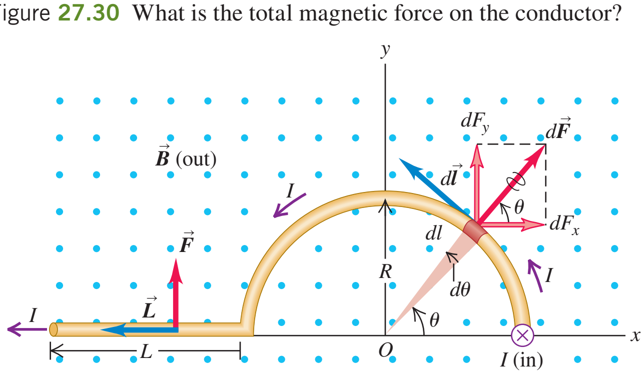
$$F_{2x} = IRB\int_0^{\pi} \cos\theta\,d\theta = 0$$

$$F_{2y} = IRB\int_0^{\pi} \sin\theta\,d\theta = 2IRB$$

Hence $\vec{F}_2 = 2IRB\hat{\jmath}$. Finally, adding the forces on all three segments, we find that the total force is in the positive y-direction:

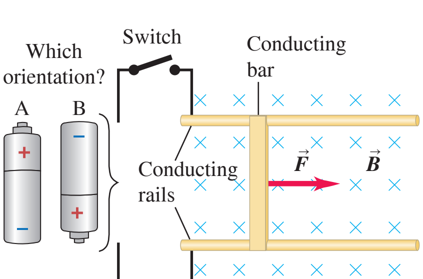
$$\vec{F} = \vec{F}_1 + \vec{F}_2 + \vec{F}_3 = 0 + 2IRB\hat{\jmath} + ILB\hat{\jmath} = IB(2R + L)\hat{\jmath}$$

Figure **27.30** What is the total magnetic force on the conductor?



EVALUATE We could have predicted from symmetry that the x-component of $\vec{F}_2$ would be zero: On the right half of the semicircle the x-component of the force is positive (to the right) and on the left half it is negative (to the left); the positive and negative contributions to the integral cancel. The result is that $\vec{F}_2$ is the force that would be exerted if we replaced the semicircle with a *straight* segment of length $2R$ along the x-axis. Do you see why?

KEYCONCEPT To find the magnetic force on a curved, current-carrying conductor in a magnetic field, first divide the conductor into infinitesimally small straight segments. Then find the force on one such segment. Finally, integrate over all segments in the conductor to find the net force.



TEST YOUR UNDERSTANDING OF SECTION 27.6 The accompanying figure shows a top view of two conducting rails on which a conducting bar can slide. A uniform magnetic field is directed perpendicular to the plane of the figure as shown. A battery is to be connected to the two rails so that when the switch is closed, current will flow through the bar and cause a magnetic force to push the bar to the right. In which orientation, A or B, should the battery be placed in the circuit?

ANSWER

I A This orientation will cause current to flow clockwise around the circuit and hence through the conducting bar in the direction from the top to the bottom of the figure. From the right-hand rule, the magnetic force $\vec{F} = I\vec{L} \times \vec{B}$ on the bar will then point to the right.

27.7 FORCE AND TORQUE ON A CURRENT LOOP

Current-carrying conductors usually form closed loops, so it is worthwhile to use the results of Section 27.6 to find the *total* magnetic force and torque on a conductor in the form of a loop. Many practical devices make use of the magnetic force or torque on a conducting loop, including loudspeakers (see Fig. 27.28) and galvanometers (see Section 26.3). Hence the results of this section are of substantial practical importance. These results will also help us understand the behavior of bar magnets described in Section 27.1.

As an example, let's look at a rectangular current loop in a uniform magnetic field. We can represent the loop as a series of straight line segments. We'll find that the total *force* on the loop is zero but that there can be a net *torque* acting on the loop, with some interesting properties.

Figure 27.31a shows a rectangular loop of wire with side lengths a and b. A line perpendicular to the plane of the loop (i.e., a *normal* to the plane) makes an angle ϕ with the direction of the magnetic field $\vec{B}$, and the loop carries a current I. The wires leading the current into and out of the loop and the source of emf are omitted to keep the diagram simple.

The force $\vec{F}$ on the right side of the loop (length a) is to the right, in the $+x$-direction as shown. On this side, $\vec{B}$ is perpendicular to the current direction, and the force on this side has magnitude

$$F = IaB \tag{27.21}$$

A force $-\vec{F}$ with the same magnitude but opposite direction acts on the opposite side of the loop, as shown in the figure.

The sides with length b make an angle $(90° - \phi)$ with the direction of $\vec{B}$. The forces on these sides are the vectors $\vec{F}'$ and $-\vec{F}'$; their magnitude F' is given by

$$F' = IbB\sin(90° - \phi) = IbB\cos\phi$$

The lines of action of both forces lie along the y-axis.

The *total* force on the loop is zero because the forces on opposite sides cancel out in pairs.

The net force on a current loop in a uniform magnetic field is zero. However, the net torque is not in general equal to zero.

(You may find it helpful to review the discussion of torque in Section 10.1.) The two forces $\vec{F}'$ and $-\vec{F}'$ in Fig. 27.31a lie along the same line and so give rise to zero net torque with

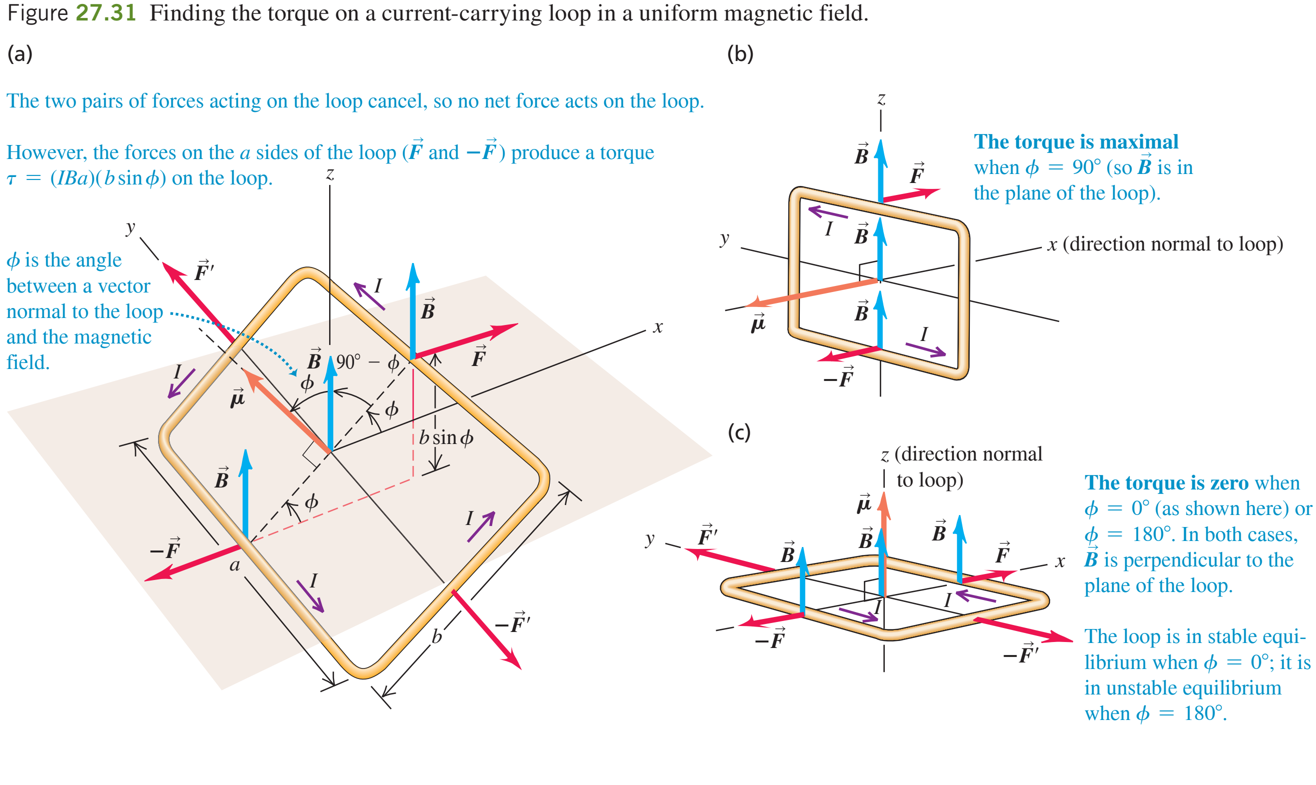
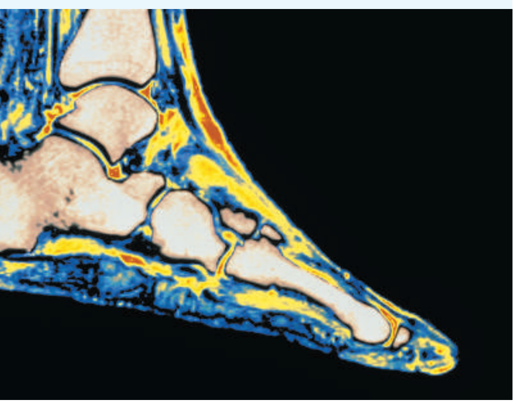


Figure **27.31** Finding the torque on a current-carrying loop in a uniform magnetic field.

BIO APPLICATION Magnetic Resonance Imaging In magnetic resonance imaging (MRI), a patient is placed in a strong magnetic field. Each hydrogen nucleus in the patient acts like a miniature current loop with a magnetic dipole moment that tends to align with the applied field. Radio waves of just the right frequency then flip these magnetic moments out of alignment. The extent to which the radio waves are absorbed is proportional to the amount of hydrogen present. This makes it possible to image details in hydrogen-rich soft tissue that cannot be seen in x-ray images. (X rays are superior to MRI for imaging bone, which is hydrogen deficient.)



respect to any point. The two forces $\vec{F}$ and $-\vec{F}$ lie along different lines, and each gives rise to a torque about the y-axis. According to the right-hand rule for determining the direction of torques, the vector torques due to $\vec{F}$ and $-\vec{F}$ are both in the $+y$-direction; hence the net vector torque $\vec{\tau}$ is in the $+y$-direction as well. The moment arm for each of these forces (equal to the perpendicular distance from the rotation axis to the line of action of the force) is $(b/2)\sin\phi$, so the torque due to each force has magnitude $F(b/2)\sin\phi$. If we use Eq. (27.21) for F, the magnitude of the net torque is

$$\tau = 2F(b/2)\sin\phi = (IBa)(b\sin\phi) \tag{27.22}$$

The torque is greatest when $\phi = 90°$, $\vec{B}$ is in the plane of the loop, and the normal to this plane is perpendicular to $\vec{B}$ (Fig. 27.31b). The torque is zero when ϕ is 0° or 180° and the normal to the loop is parallel or antiparallel to the field (Fig. 27.31c). The value $\phi = 0°$ is a stable equilibrium position because the torque is zero there, and when the loop is rotated slightly from this position, the resulting torque tends to rotate it back toward $\phi = 0°$. The position $\phi = 180°$ is an *unstable* equilibrium position; if displaced slightly from this position, the loop tends to move farther away from $\phi = 180°$. Figure 27.31 shows rotation about the y-axis, but because the net force on the loop is zero, Eq. (27.22) for the torque is valid for *any* choice of axis. The torque always tends to rotate the loop in the direction of *decreasing* ϕ—that is, toward the stable equilibrium position $\phi = 0°$.

The area A of the loop is equal to ab, so we can rewrite Eq. (27.22) as

Magnitude of magnetic torque on a current loop

$$\tau = IBA\sin\phi \tag{27.23}$$

Current; Magnetic-field magnitude; Area of loop; Angle between normal to loop plane and field direction

The product IA is called the **magnetic dipole moment** or **magnetic moment** of the loop, for which we use the symbol μ (the Greek letter mu):

$$\mu = IA \tag{27.24}$$

It is analogous to the electric dipole moment introduced in Section 21.7. In terms of μ, the magnitude of the torque on a current loop is

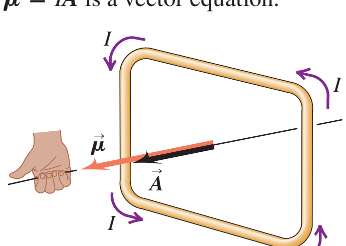
$$\tau = \mu B\sin\phi \tag{27.25}$$

where ϕ is the angle between the normal to the loop (the direction of the vector area $\vec{A}$) and $\vec{B}$. A current loop, or any other object that experiences a magnetic torque given by Eq. (27.25), is also called a **magnetic dipole.**

Magnetic Torque: Vector Form

We can also define a vector magnetic moment $\vec{\mu}$ with magnitude IA: This is shown in Fig. 27.31. The direction of $\vec{\mu}$ is defined to be perpendicular to the plane of the loop, with a sense determined by a right-hand rule, as shown in **Fig. 27.32**. Wrap the fingers of your right hand around the perimeter of the loop in the direction of the current. Then extend your thumb so that it is perpendicular to the plane of the loop; its direction is the direction of $\vec{\mu}$ (and of the vector area $\vec{A}$ of the loop). The torque is greatest when $\vec{\mu}$ and $\vec{B}$ are perpendicular and is zero when they are parallel or antiparallel. In the stable equilibrium position, $\vec{\mu}$ and $\vec{B}$ are parallel.

Figure 27.32 The right-hand rule determines the direction of the magnetic moment of a current-carrying loop. This is also the direction of the loop's area vector $\vec{A}$; $\vec{\mu} = I\vec{A}$ is a vector equation.



Finally, we can express this interaction in terms of the torque vector $\vec{\tau}$, which we used for *electric*-dipole interactions in Section 21.7. From Eq. (27.25) the magnitude of $\vec{\tau}$ is equal to the magnitude of $\vec{\mu}\times\vec{B}$, and reference to Fig. 27.31 shows that the directions are also the same. So we have

Vector magnetic torque on a current loop

$$\vec{\tau} = \vec{\mu}\times\vec{B} \tag{27.26}$$

Magnetic dipole moment; Magnetic field

This result is directly analogous to the result we found in Section 21.7 for the torque exerted by an *electric* field $\vec{E}$ on an *electric* dipole with dipole moment $\vec{p}$.

Potential Energy for a Magnetic Dipole

When a magnetic dipole changes orientation in a magnetic field, the field does work on it. In an infinitesimal angular displacement $d\phi$, the work dW is given by $\tau\, d\phi$, and there is a corresponding change in potential energy. As the above discussion suggests, the potential energy U is least when $\vec{\mu}$ and $\vec{B}$ are parallel and greatest when they are antiparallel. To find an expression for U as a function of orientation, note that the torque on an *electric* dipole in an *electric* field is $\vec{\tau} = \vec{p} \times \vec{E}$; we found in Section 21.7 that the corresponding potential energy is $U = -\vec{p} \cdot \vec{E}$. The torque on a *magnetic* dipole in a *magnetic* field is $\vec{\tau} = \vec{\mu} \times \vec{B}$, so we can conclude immediately that

Potential energy for a magnetic dipole in a magnetic field

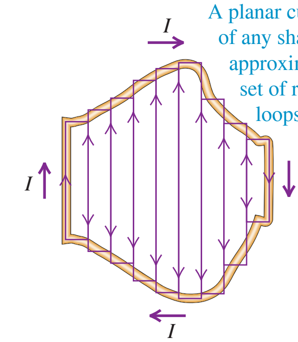
$$U = -\underbrace{\vec{\mu}}_{\text{Magnetic dipole moment}} \cdot \underbrace{\vec{B}}_{\text{Magnetic field}} = -\mu B \cos \underbrace{\phi}_{\text{Angle between } \vec{\mu} \text{ and } \vec{B}} \qquad (27.27)$$

With this definition, U is zero when the magnetic dipole moment is perpendicular to the magnetic field ($\phi = 90°$); then $U = -\vec{\mu} \cdot \vec{B} = -\mu B \cos 90° = 0$.

Magnetic Torque: Loops and Coils

Although we have derived Eqs. (27.21) through (27.27) for a rectangular current loop, all these relationships are valid for a plane loop of any shape at all. Any planar loop may be approximated as closely as we wish by a very large number of rectangular loops, as shown in **Fig. 27.33**. If these loops all carry equal currents in the same clockwise sense, then the forces and torques on the sides of two loops adjacent to each other cancel, and the only forces and torques that do not cancel are due to currents around the boundary. Thus all the above relationships are valid for a plane current loop of any shape, with the magnetic moment $\vec{\mu} = I\vec{A}$.

Figure **27.33** The collection of rectangles exactly matches the irregular plane loop in the limit as the number of rectangles approaches infinity and the width of each rectangle approaches zero.



We can also generalize this whole formulation to a coil consisting of N planar loops close together; the effect is simply to multiply each force, the magnetic moment, the torque, and the potential energy by a factor of N.

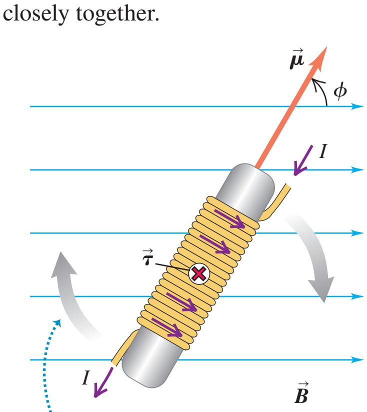
An arrangement of particular interest is the **solenoid,** a helical winding of wire, such as a coil wound on a circular cylinder (**Fig. 27.34**). If the windings are closely spaced, the solenoid can be approximated by a number of circular loops lying in planes at right angles to its long axis. The total torque on a solenoid in a magnetic field is simply the sum of the torques on the individual turns. For a solenoid with N turns in a uniform field B, the magnetic moment is $\mu = NIA$ and

$$\tau = NIAB \sin \phi \qquad (27.28)$$

where ϕ is the angle between the axis of the solenoid and the direction of the field. The magnetic-moment vector $\vec{\mu}$ is along the solenoid axis. The torque is greatest when the solenoid axis is perpendicular to the magnetic field and zero when they are parallel. The effect of this torque is to tend to rotate the solenoid into a position where its axis is parallel to the field. Solenoids are also useful as *sources* of magnetic field, as we'll discuss in Chapter 28.

The d'Arsonval galvanometer, described in Section 26.3, makes use of a magnetic torque on a coil carrying a current. As Fig. 26.14 shows, the magnetic field is not uniform but is *radial,* so the side thrusts on the coil are always perpendicular to its plane. Thus the angle ϕ in Eq. (27.28) is always 90°, and the magnetic torque is directly proportional to the current, no matter what the orientation of the coil. A restoring torque proportional to the angular displacement of the coil is provided by two hairsprings, which also serve as current leads to the coil. When current is supplied to the coil, it rotates along with its attached pointer until the restoring spring torque just balances the magnetic torque. Thus the pointer deflection is proportional to the current.

Figure **27.34** The torque $\vec{\tau} = \vec{\mu} \times \vec{B}$ on this solenoid in a uniform magnetic field is directed straight into the page. An actual solenoid has many more turns, wrapped closely together.

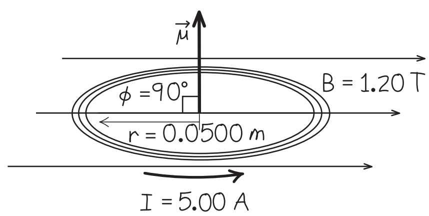


EXAMPLE 27.9 Magnetic torque on a circular coil

A circular coil 0.0500 m in radius, with 30 turns of wire, lies in a horizontal plane. It carries a counterclockwise (as viewed from above) current of 5.00 A. The coil is in a uniform 1.20 T magnetic field directed toward the right. Find the magnitudes of the magnetic moment and the torque on the coil.

IDENTIFY and SET UP This problem uses the definition of magnetic moment and the expression for the torque on a magnetic dipole in a magnetic field. **Figure 27.35** shows the situation. Equation (27.24) gives the magnitude μ of the magnetic moment of a single turn of wire; for N turns, the magnetic moment is N times greater. Equation (27.25) gives the magnitude τ of the torque.

Figure 27.35 Our sketch for this problem.



EXECUTE The area of the coil is $A = \pi r^2$. From Eq. (27.24), the total magnetic moment of all 30 turns is

$$\mu_{\text{total}} = NIA = 30(5.00\text{ A})\pi(0.0500\text{ m})^2 = 1.18\text{ A}\cdot\text{m}^2$$

The angle ϕ between the direction of $\vec{B}$ and the direction of $\vec{\mu}$ (which is along the normal to the plane of the coil) is 90°. From Eq. (27.25), the torque on the coil is

$$\begin{aligned}\tau &= \mu_{\text{total}} B \sin\phi = (1.18\text{ A}\cdot\text{m}^2)(1.20\text{ T})(\sin 90°)\\ &= 1.41\text{ N}\cdot\text{m}\end{aligned}$$

EVALUATE The torque tends to rotate the right side of the coil down and the left side up, into a position where the normal to its plane is parallel to $\vec{B}$.

KEYCONCEPT A current-carrying coil in a magnetic field $\vec{B}$ experiences a magnetic torque. The torque tends to orient the coil so its magnetic dipole moment $\vec{\mu}$ (a vector perpendicular to the plane of the coil, with magnitude equal to the product of the number of turns, the current, and the coil area) is in the same direction as $\vec{B}$.

EXAMPLE 27.10 Potential energy for a coil in a magnetic field

If the coil in Example 27.9 rotates from its initial orientation to one in which its magnetic moment $\vec{\mu}$ is parallel to $\vec{B}$, what is the change in potential energy?

IDENTIFY and SET UP Equation (27.27) gives the potential energy for each orientation. The initial position is as shown in Fig. 27.35, with $\phi_1 = 90°$. In the final orientation, the coil has been rotated 90° clockwise so that $\vec{\mu}$ and $\vec{B}$ are parallel, so the angle between these vectors is $\phi_2 = 0$.

EXECUTE From Eq. (27.27), the potential energy change is

$$\begin{aligned}\Delta U &= U_2 - U_1 = -\mu B\cos\phi_2 - (-\mu B\cos\phi_1)\\ &= -\mu B(\cos\phi_2 - \cos\phi_1)\\ &= -(1.18\text{ A}\cdot\text{m}^2)(1.20\text{ T})(\cos 0° - \cos 90°) = -1.41\text{ J}\end{aligned}$$

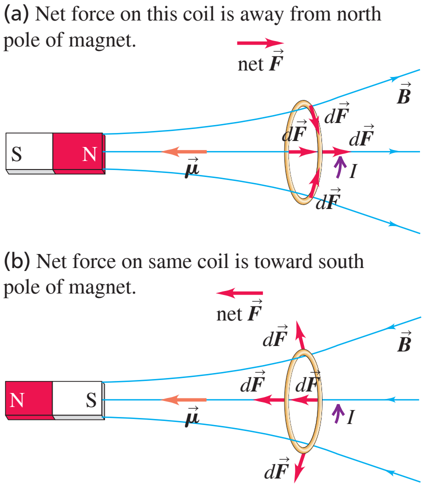
EVALUATE The potential energy decreases because the rotation is in the direction of the magnetic torque that we found in Example 27.9.

KEYCONCEPT The potential energy of a current-carrying coil in a magnetic field $\vec{B}$ depends on the magnitudes and directions of $\vec{B}$ and of the coil's magnetic dipole moment $\vec{\mu}$. The potential energy is minimum when $\vec{\mu}$ and $\vec{B}$ are in the same direction.

Figure 27.36 Forces on current loops in a nonuniform $\vec{B}$ field. In each case the axis of the bar magnet is perpendicular to the plane of the loop and passes through the center of the loop.

(a) Net force on this coil is away from north pole of magnet.

(b) Net force on same coil is toward south pole of magnet.



Magnetic Dipole in a Nonuniform Magnetic Field

We have seen that a current loop (that is, a magnetic dipole) experiences zero net force in a uniform magnetic field. **Figure 27.36** shows two current loops in the *nonuniform* $\vec{B}$ field of a bar magnet; in both cases the net force on the loop is *not* zero. In Fig. 27.36a the magnetic moment $\vec{\mu}$ is in the direction opposite to the field, and the force $d\vec{F} = I\,d\vec{l} \times \vec{B}$ on a segment of the loop has both a radial component and a component to the right. When these forces are summed to find the net force $\vec{F}$ on the loop, the radial components cancel so that the net force is to the right, away from the magnet. Note that in this case the force is toward the region where the field lines are farther apart and the field magnitude B is less. The polarity of the bar magnet is reversed in Fig. 27.36b, so $\vec{\mu}$ and $\vec{B}$ are parallel; now the net force on the loop is to the left, toward the region of greater field magnitude near the magnet. Later in this section we'll use these observations to explain why bar magnets can pick up unmagnetized iron objects.

Magnetic Dipoles and How Magnets Work

The behavior of a solenoid in a magnetic field (see Fig. 27.34) resembles that of a bar magnet or compass needle; if free to turn, both the solenoid and the magnet orient themselves with their axes parallel to the $\vec{B}$ field. In both cases this is due to the interaction of moving electric charges with a magnetic field; the difference is that in a bar magnet the motion of charge occurs on the microscopic scale of the atom.

Think of an electron as being like a spinning ball of charge. In this analogy the circulation of charge around the spin axis is like a current loop, and so the electron has a net magnetic moment. (This analogy, while helpful, is inexact; an electron isn't really a spinning sphere. A full explanation of the origin of an electron's magnetic moment involves quantum mechanics, which is beyond our scope here.) In an iron atom a substantial fraction of the electron magnetic moments align with each other, and the atom has a nonzero magnetic moment. (By contrast, the atoms of most elements have little or no net magnetic moment.) In an unmagnetized piece of iron there is no overall alignment of the magnetic moments of the atoms; their vector sum is zero, and the net magnetic moment is zero (**Fig. 27.37a**). But in an iron bar magnet the magnetic moments of many of the atoms are parallel, and there is a substantial net magnetic moment $\vec{\mu}$ (Fig. 27.37b). If the magnet is placed in a magnetic field $\vec{B}$, the field exerts a torque given by Eq. (27.26) that tends to align $\vec{\mu}$ with $\vec{B}$ (Fig. 27.37c). A bar magnet tends to align with a $\vec{B}$ field so that a line from the south pole to the north pole of the magnet is in the direction of $\vec{B}$; hence the real significance of a magnet's north and south poles is that they represent the head and tail, respectively, of the magnet's dipole moment $\vec{\mu}$.

The torque experienced by a current loop in a magnetic field also explains how an unmagnetized iron object like that in Fig. 27.37a becomes magnetized. If an unmagnetized iron paper clip is placed next to a powerful magnet, the magnetic moments of the paper clip's atoms tend to align with the $\vec{B}$ field of the magnet. When the paper clip is removed, its atomic dipoles tend to remain aligned, and the paper clip has a net magnetic moment. The paper clip can be demagnetized by being dropped on the floor or heated; the added internal energy jostles and re-randomizes the atomic dipoles.

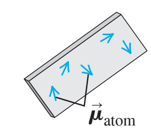
The magnetic-dipole picture of a bar magnet explains the attractive and repulsive forces between bar magnets shown in Fig. 27.1. The magnetic moment $\vec{\mu}$ of a bar magnet points from its south pole to its north pole, so the current loops in Figs. 27.36a and 27.36b are both equivalent to a magnet with its north pole on the left. Hence the situation in Fig. 27.36a is equivalent to two bar magnets with their north poles next to each other; the resultant force is repulsive, as in Fig. 27.1b. In Fig. 27.36b we again have the equivalent of two bar magnets end to end, but with the south pole of the left-hand magnet next to the north pole of the right-hand magnet. The resultant force is attractive, as in Fig. 27.1a.

Finally, we can explain how a magnet can attract an unmagnetized iron object (see Fig. 27.2). It's a two-step process. First, the atomic magnetic moments of the iron tend to align with the $\vec{B}$ field of the magnet, so the iron acquires a net magnetic dipole moment $\vec{\mu}$ parallel to the field. Second, the nonuniform field of the magnet attracts the magnetic dipole. **Figure 27.38a** shows an example. The north pole of the magnet is closer to the nail (which contains iron), and the magnetic dipole produced in the nail is equivalent to a loop with a current that circulates in a direction opposite to that shown in Fig. 27.36a. Hence the net magnetic force on the nail is opposite to the force on the loop in Fig. 27.36a, and the nail is attracted toward the magnet. Changing the polarity of the magnet, as in Fig. 27.38b, reverses the directions of both $\vec{B}$ and $\vec{\mu}$. The situation is now equivalent to that shown in Fig. 27.36b; like the loop in that figure, the nail is attracted toward the magnet. Hence a previously unmagnetized object containing iron is attracted to *either* pole of a magnet. By contrast, objects made of brass, aluminum, or wood hardly respond at all to a magnet; the atomic magnetic dipoles of these materials, if present at all, have less tendency to align with an external field.

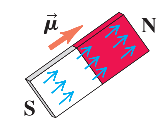
Our discussion of how magnets and pieces of iron interact has just scratched the surface of a diverse subject known as *magnetic properties of materials*. We'll discuss these properties in more depth in Section 28.8.

Figure **27.37** (a) An unmagnetized piece of iron. (Only a few representative atomic moments are shown.) (b) A magnetized piece of iron (bar magnet). The net magnetic moment of the bar magnet points from its south pole to its north pole. (c) A bar magnet in a magnetic field.

(a) Unmagnetized iron: magnetic moments are oriented randomly.



(b) In a bar magnet, the magnetic moments are aligned.



(c) A magnetic field creates a torque on the bar magnet that tends to align its dipole moment with the $\vec{B}$ field.

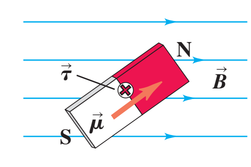
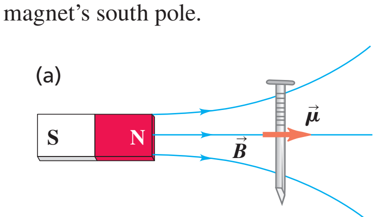
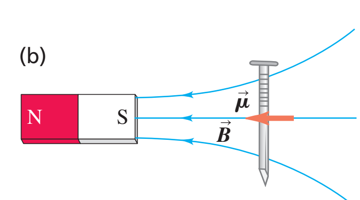


Figure **27.38** A bar magnet attracts an unmagnetized iron nail in two steps. First, the $\vec{B}$ field of the bar magnet gives rise to a net magnetic moment in the nail. Second, because the field of the bar magnet is not uniform, this magnetic dipole is attracted toward the magnet. The attraction is the same whether the nail is closer to (a) the magnet's north pole or (b) the magnet's south pole.

(a)



(b)



TEST YOUR UNDERSTANDING OF SECTION 27.7 Figure 27.13c depicts the magnetic field lines due to a circular current-carrying loop. (a) What is the direction of the magnetic moment of this loop? (b) Which side of the loop is equivalent to the north pole of a magnet, and which side is equivalent to the south pole?

ANSWER

❙ (a) **to the right;** (b) **north pole on the right, south pole on the left** If you wrap the fingers of your right hand around the coil in the direction of the current, your right thumb points to the right (perpendicular to the plane of the coil). This is the direction of the magnetic moment $\vec{\mu}$. The magnetic moment points from the south pole to the north pole, so the right side of the loop is equivalent to a north pole and the left side is equivalent to a south pole.

27.8 THE DIRECT-CURRENT MOTOR

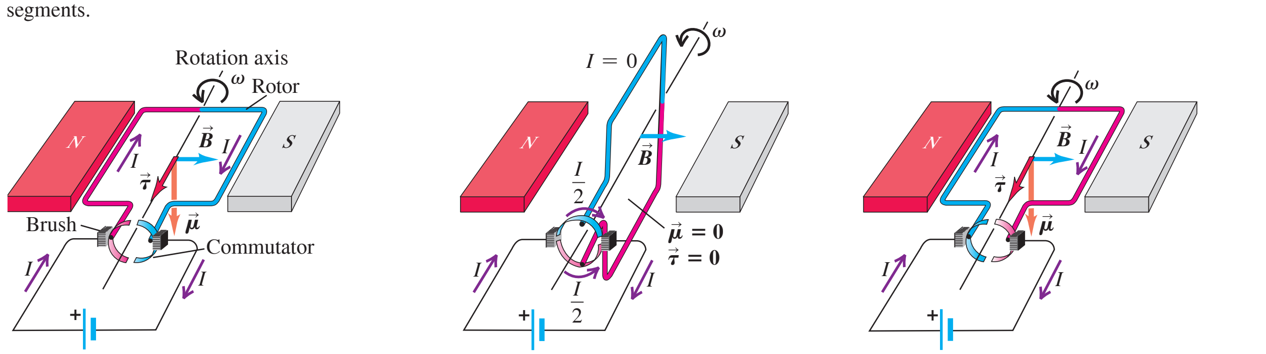
Electric motors play an important role in contemporary society. In a motor a magnetic torque acts on a current-carrying conductor, and electric energy is converted to mechanical energy. As an example, let's look at a simple type of direct-current (dc) motor, shown in **Fig. 27.39**.

The moving part of the motor is the *rotor,* a length of wire formed into an open-ended loop and free to rotate about an axis. The ends of the rotor wires are attached to circular conducting segments that form a *commutator.* In Fig. 27.39a, each of the two commutator segments makes contact with one of the terminals, or *brushes,* of an external circuit that includes a source of emf. This causes a current to flow into the rotor on one side, shown in red, and out of the rotor on the other side, shown in blue. Hence the rotor is a current loop with a magnetic moment $\vec{\mu}$. The rotor lies between opposing poles of a permanent magnet, so there is a magnetic field $\vec{B}$ that exerts a torque $\vec{\tau} = \vec{\mu} \times \vec{B}$ on the rotor. For the rotor orientation shown in Fig. 27.39a the torque causes the rotor to turn counterclockwise, in the direction that will align $\vec{\mu}$ with $\vec{B}$.

In Fig. 27.39b the rotor has rotated by 90° from its orientation in Fig. 27.39a. If the current through the rotor were constant, the rotor would now be in its equilibrium orientation; it would simply oscillate around this orientation. But here's where the commutator comes into play; each brush is now in contact with *both* segments of the commutator. There is no potential difference between the commutators, so at this instant no current flows through the rotor, and the magnetic moment is zero. The rotor continues to rotate counterclockwise because of its inertia, and current again flows through the rotor as in Fig. 27.39c. But now current enters on the *blue* side of the rotor and exits on the *red* side, just the opposite of the situation in Fig. 27.39a. While the direction of the current has reversed with respect

Figure **27.39** Schematic diagram of a simple dc motor. The rotor is a wire loop that is free to rotate about an axis; the rotor ends are attached to the two curved conductors that form the commutator. (The rotor halves are colored red and blue for clarity.) The commutator segments are insulated from one another.

(a) Brushes are aligned with commutator segments. (b) Rotor has turned 90°. (c) Rotor has turned 180°.



- Current flows into the red side of the rotor and out of the blue side.
- Therefore the magnetic torque causes the rotor to spin counterclockwise.

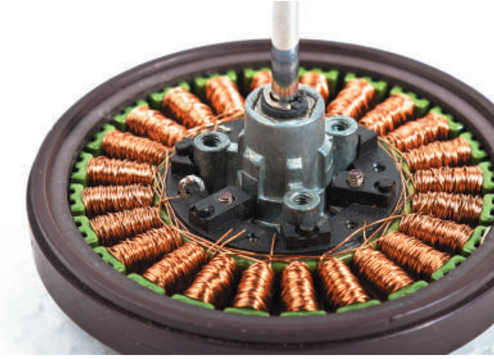
- Each brush is in contact with both commutator segments, so the current bypasses the rotor altogether.
- No magnetic torque acts on the rotor.

- The brushes are again aligned with commutator segments. This time the current flows into the blue side of the rotor and out of the red side.
- Therefore the magnetic torque again causes the rotor to spin counterclockwise.

to the rotor, the rotor itself has rotated 180° and the magnetic moment $\vec{\boldsymbol{\mu}}$ is in the same direction with respect to the magnetic field. Hence the magnetic torque $\vec{\boldsymbol{\tau}}$ is in the same direction in Fig. 27.39c as in Fig. 27.39a. Thanks to the commutator, the current reverses after every 180° of rotation, so the torque is always in the direction to rotate the rotor counterclockwise. When the motor has come "up to speed," the average magnetic torque is just balanced by an opposing torque due to air resistance, friction in the rotor bearings, and friction between the commutator and brushes.

The simple motor shown in Fig. 27.39 has only a single turn of wire in its rotor. In practical motors, the rotor has many turns; this increases the magnetic moment and the torque so that the motor can spin larger loads. The torque can also be increased by using a stronger magnetic field, which is why many motor designs use electromagnets instead of a permanent magnet. Another drawback of the simple design in Fig. 27.39 is that the magnitude of the torque rises and falls as the rotor spins. This can be remedied by having the rotor include several independent coils of wire oriented at different angles (**Fig. 27.40**).

Figure **27.40** This electric motor has multiple current-carrying coils. They interact with permanent magnets on the rotor (not shown) to make the rotor spin around the motor's shaft. (This design is the reverse of the design in Fig. 27.39, in which the permanent magnets are stationary and the coil rotates.) Because there are multiple coils, the magnetic torque is very nearly constant and the rotor spins at a constant rate. Motors of this kind are used in drones, electric cars, and many other applications.



Power for Electric Motors

Because a motor converts electric energy to mechanical energy or work, it requires electric energy input. If the potential difference between its terminals is V_{ab} and the current is I, then the power input is $P = V_{ab}I$. Even if the motor coils have negligible resistance, there must be a potential difference between the terminals if P is to be different from zero. This potential difference results principally from magnetic forces exerted on the currents in the conductors of the rotor as they rotate through the magnetic field. The associated electromotive force $\mathcal{E}$ is called an *induced* emf; it is also called a *back* emf because its sense is opposite to that of the current. In Chapter 29 we'll study induced emfs resulting from motion of conductors in magnetic fields.

In a *series* motor the rotor is connected in series with the electromagnet that produces the magnetic field; in a *shunt* motor they are connected in parallel. In a series motor with internal resistance r, V_{ab} is greater than $\mathcal{E}$, and the difference is the potential drop Ir across the internal resistance. That is,

$$V_{ab} = \mathcal{E} + Ir \tag{27.29}$$

Because the magnetic force is proportional to velocity, $\mathcal{E}$ is *not* constant but is proportional to the speed of rotation of the rotor.

EXAMPLE 27.11 A series dc motor

A dc motor with its rotor and field coils connected in series has an internal resistance of 2.00 Ω. When running at full load on a 120 V line, it draws a 4.00 A current. (a) What is the emf in the rotor? (b) What is the power delivered to the motor? (c) What is the rate of dissipation of energy in the internal resistance? (d) What is the mechanical power developed? (e) What is the motor's efficiency? (f) What happens if the machine being driven by the motor jams, so that the rotor suddenly stops turning?

IDENTIFY and SET UP This problem uses the ideas of power and potential drop in a series dc motor. We are given the internal resistance $r = 2.00\ \Omega$, the voltage $V_{ab} = 120$ V across the motor, and the current $I = 4.00$ A through the motor. We use Eq. (27.29) to determine the emf $\mathcal{E}$ from these quantities. The power delivered to the motor is $V_{ab}I$, the rate of energy dissipation is I^2r, and the power output by the motor is the difference between the power input and the power dissipated. The efficiency e is the ratio of mechanical power output to electric power input.

EXECUTE (a) From Eq. (27.29), $V_{ab} = \mathcal{E} + Ir$, we have

$$120\text{ V} = \mathcal{E} + (4.00\text{ A})(2.00\ \Omega) \qquad \text{and so} \qquad \mathcal{E} = 112\text{ V}$$

(b) The power delivered to the motor from the source is

$$P_{\text{input}} = V_{ab}I = (120\text{ V})(4.00\text{ A}) = 480\text{ W}$$

(c) The power dissipated in the resistance r is

$$P_{\text{dissipated}} = I^2r = (4.00\text{ A})^2(2.00\ \Omega) = 32\text{ W}$$

(d) The mechanical power output is the electric power input minus the rate of dissipation of energy in the motor's resistance (assuming that there are no other power losses):

$$P_{\text{output}} = P_{\text{input}} - P_{\text{dissipated}} = 480\text{ W} - 32\text{ W} = 448\text{ W}$$

(e) The efficiency e is the ratio of mechanical power output to electric power input:

$$e = \frac{P_{\text{output}}}{P_{\text{input}}} = \frac{448\text{ W}}{480\text{ W}} = 0.93 = 93\%$$

(f) With the rotor stalled, the back emf $\mathcal{E}$ (which is proportional to rotor speed) goes to zero. From Eq. (27.29) the current becomes

$$I = \frac{V_{ab}}{r} = \frac{120\text{ V}}{2.00\ \Omega} = 60\text{ A}$$

and the power dissipated in the resistance r becomes

$$P_{\text{dissipated}} = I^2r = (60\text{ A})^2(2.00\ \Omega) = 7200\text{ W}$$

Continued

EVALUATE If this massive overload doesn't blow a fuse or trip a circuit breaker, the coils will quickly melt. When the motor is first turned on, there's a momentary surge of current until the motor picks up speed. This surge causes greater-than-usual voltage drops ($V = IR$) in the power lines supplying the current. Similar effects are responsible for the momentary dimming of lights that can occur in a house when an air conditioner or dishwasher motor starts.

KEYCONCEPT When an electric motor is in operation, a large emf is induced in the rotor coils. To find the mechanical power developed by the motor, multiply this emf by the current in the coils.

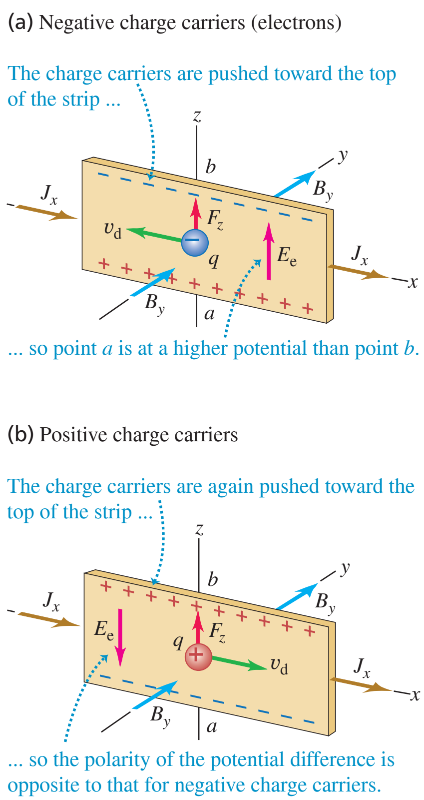
TEST YOUR UNDERSTANDING OF SECTION 27.8 In the circuit shown in Fig. 27.39, you add a switch in series with the source of emf so that the current can be turned on and off. When you close the switch and allow current to flow, will the rotor begin to turn no matter what its original orientation?

ANSWER

no The rotor will not begin to turn when the switch is closed if the rotor is initially oriented as shown in Fig. 27.39b. In this case there is no current through the rotor and hence no magnetic torque. This situation can be remedied by using multiple rotor coils oriented at different angles around the rotation axis. With this arrangement, there is always a magnetic torque no matter what the orientation.

27.9 THE HALL EFFECT

Figure 27.41 Forces on charge carriers in a conductor in a magnetic field.



The reality of the forces acting on the moving charges in a conductor in a magnetic field is strikingly demonstrated by the *Hall effect,* an effect analogous to the transverse deflection of an electron beam in a magnetic field in vacuum. (The effect was discovered by the American physicist Edwin Hall in 1879 while he was still a graduate student.) To describe this effect, let's consider a conductor in the form of a flat strip, as shown in **Fig. 27.41**. The current is in the direction of the $+x$-axis and there is a uniform magnetic field $\vec{B}$ perpendicular to the plane of the strip, in the $+y$-direction. The drift velocity of the moving charges (charge magnitude $|q|$) has magnitude v_d. Figure 27.41a shows the case of negative charges, such as electrons in a metal, and Fig. 27.41b shows positive charges. In both cases the magnetic force is upward, just as the magnetic force on a conductor is the same whether the moving charges are positive or negative. In either case a moving charge is driven toward the *upper* edge of the strip by the magnetic force $F_z = |q|v_dB$.

If the charge carriers are electrons, as in Fig. 27.41a, an excess negative charge accumulates at the upper edge of the strip, leaving an excess positive charge at its lower edge. This accumulation continues until the resulting transverse electrostatic field $\vec{E}_e$ becomes large enough to cause a force (magnitude $|q|E_e$) that is equal and opposite to the magnetic force (magnitude $|q|v_dB$). After that, there is no longer any net transverse force to deflect the moving charges. This electric field causes a transverse potential difference between opposite edges of the strip, called the *Hall voltage* or the *Hall emf.* The polarity depends on whether the moving charges are positive or negative. Experiments show that for metals the upper edge of the strip in Fig. 27.41a *does* become negatively charged, showing that the charge carriers in a metal are indeed negative electrons.

However, if the charge carriers are *positive,* as in Fig. 27.41b, then *positive* charge accumulates at the upper edge, and the potential difference is *opposite* to the situation with negative charges. Soon after the discovery of the Hall effect in 1879, it was observed that some materials, particularly some *semiconductors,* show a Hall emf opposite to that of the metals, as if their charge carriers were positively charged. We now know that these materials conduct by a process known as *hole conduction.* Within such a material there are locations, called *holes,* that would normally be occupied by an electron but are actually empty. A missing negative charge is equivalent to a positive charge. When an electron moves in one direction to fill a hole, it leaves another hole behind it. The hole migrates in the direction opposite to that of the electron.

In terms of the coordinate axes in Fig. 27.41b, the electrostatic field $\vec{E}_e$ for the positive-q case is in the $-z$-direction; its z-component E_z is negative. The magnetic field is in the $+y$-direction, and we write it as B_y. The magnetic force (in the $+z$-direction) is qv_dB_y. The current density J_x is in the $+x$-direction. In the steady state, when the forces qE_z and qv_dB_y sum to zero,

$$qE_z + qv_dB_y = 0 \qquad \text{or} \qquad E_z = -v_dB_y$$

This confirms that when q is positive, E_z is negative. From Eq. (25.4),

$$J_x = nqv_d$$

Eliminating v_d between these equations, we find

Hall effect: $$nq = \frac{-J_x B_y}{E_z} \tag{27.30}$$

Concentration of mobile charge carriers; Charge per carrier; Current density; Magnetic field; Electrostatic field in conductor

Note that this result (as well as the entire derivation) is valid for both positive and negative q. When q is negative, E_z is positive, and conversely.

We can measure J_x, B_y, and E_z, so we can compute the product nq. In both metals and semiconductors, q is equal in magnitude to the electron charge, so the Hall effect permits a direct measurement of n, the concentration of current-carrying charges in the material. The *sign* of the charges is determined by the polarity of the Hall emf, as we have described.

The Hall effect can also be used for a direct measurement of electron drift speed v_d in metals. As we saw in Chapter 25, these speeds are very small, often of the order of 1 mm/s or less. If we move the entire conductor in the opposite direction to the current with a speed equal to the drift speed, then the electrons are at rest with respect to the magnetic field, and the Hall emf disappears. Thus the conductor speed needed to make the Hall emf vanish is equal to the drift speed.

BIO APPLICATION Exercise Machines and the Hall Effect This athlete is making a flywheel spin inside a rowing machine. A gear attached to the flywheel has magnetic teeth, and every time a tooth on the spinning gear passes a current-carrying sensor, the tooth's magnetic field deflects the current and produces a Hall voltage in the sensor. The harder the athlete works, the faster the flywheel spins and the more voltage pulses there are per minute. The voltage information is passed to the rowing machine's display and tells the athlete her power output.

EXAMPLE 27.12 A Hall-effect measurement

You place a strip of copper, 2.0 mm thick and 1.50 cm wide, in a uniform 0.40 T magnetic field as shown in Fig. 27.41a. When you run a 75 A current in the $+x$-direction, you find that the potential at the bottom of the slab is 0.81 μV higher than at the top. From this measurement, determine the concentration of mobile electrons in copper.

IDENTIFY and SET UP This problem describes a Hall-effect experiment. We use Eq. (27.30) to determine the mobile electron concentration n.

EXECUTE First we find the current density J_x and the electric field E_z:

$$J_x = \frac{I}{A} = \frac{75\text{ A}}{(2.0\times10^{-3}\text{ m})(1.50\times10^{-2}\text{ m})} = 2.5\times10^{6}\text{ A/m}^2$$

$$E_z = \frac{V}{d} = \frac{0.81\times10^{-6}\text{ V}}{1.5\times10^{-2}\text{ m}} = 5.4\times10^{-5}\text{ V/m}$$

Then, from Eq. (27.30),

$$n = \frac{-J_xB_y}{qE_z} = \frac{-(2.5\times10^{6}\text{ A/m}^2)(0.40\text{ T})}{(-1.60\times10^{-19}\text{ C})(5.4\times10^{-5}\text{ V/m})}$$
$$= 11.6\times10^{28}\text{ m}^{-3}$$

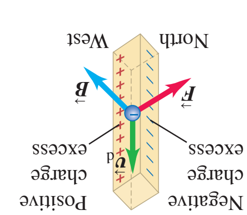
EVALUATE The actual value of n for copper is $8.5\times10^{28}\text{ m}^{-3}$. The difference shows that our simple model of the Hall effect, which ignores quantum effects and electron interactions with the ions, must be used with caution. This example also shows that with good conductors, the Hall emf is very small even with large current densities. In practice, Hall-effect devices for magnetic-field measurements use semiconductor materials, for which moderate current densities give much larger Hall emfs.

KEYCONCEPT In the Hall effect, a current in a conductor flows in a direction perpendicular to a magnetic field $\vec{B}$. The magnetic force on the current produces a charge separation in the conductor. This separation is perpendicular to both the current and $\vec{B}$, and produces a transverse electric field and potential difference in the conductor.

TEST YOUR UNDERSTANDING OF SECTION 27.9 A copper wire of square cross section is oriented vertically. The four sides of the wire face north, south, east, and west. There is a uniform magnetic field directed from east to west, and the wire carries current downward. Which side of the wire is at the highest electric potential? (i) North side; (ii) south side; (iii) east side; (iv) west side.

ANSWER

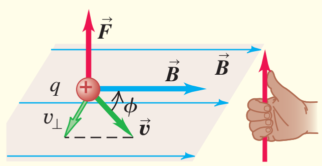
(ii) The mobile charge carriers in copper are negatively charged electrons, which move upward through the wire to give a downward current. From the right-hand rule, the force on a positively charged particle moving upward in a westward-pointing magnetic field would be to the south; hence the force on a negatively charged particle is to the north. The result is an excess of negative charge on the north side of the wire, leaving an excess of positive charge—and hence a higher electric potential—on the south side.



CHAPTER 27 SUMMARY

Magnetic forces: Magnetic interactions are fundamentally interactions between moving charged particles. These interactions are described by the vector magnetic field, denoted by $\vec{B}$. A particle with charge q moving with velocity $\vec{v}$ in a magnetic field $\vec{B}$ experiences a force $\vec{F}$ that is perpendicular to both $\vec{v}$ and $\vec{B}$. The SI unit of magnetic field is the tesla $(1\ \text{T} = 1\ \text{N/A}\cdot\text{m})$. (See Example 27.1.)

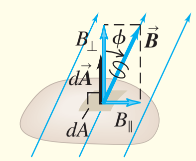
$$\vec{F} = q\vec{v} \times \vec{B} \qquad (27.2)$$



Magnetic field lines and flux: A magnetic field can be represented graphically by magnetic field lines. At each point a magnetic field line is tangent to the direction of $\vec{B}$ at that point. Where field lines are close together, the field magnitude is large, and vice versa. Magnetic flux Φ_B through an area is defined in an analogous way to electric flux; its SI unit is the weber $(1\ \text{Wb} = 1\ \text{T}\cdot\text{m}^2)$. The net magnetic flux through any closed surface is zero (Gauss's law for magnetism), so magnetic field lines always close on themselves. (See Example 27.2.)

$$\Phi_B = \int B\cos\phi\, dA = \int B_\perp\, dA = \int \vec{B}\cdot d\vec{A} \qquad (27.6)$$

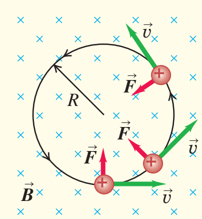
$$\oint \vec{B}\cdot d\vec{A} = 0 \quad \text{(closed surface)} \qquad (27.8)$$



Motion in a magnetic field: The magnetic force is always perpendicular to $\vec{v}$; a particle moving under the action of a magnetic field alone moves with constant speed. In a uniform field, a particle with initial velocity perpendicular to the field moves in a circle with radius R that depends on the magnetic field strength B and the particle mass m, speed v, and charge q. (See Examples 27.3 and 27.4.)

Crossed electric and magnetic fields can be used as a velocity selector. The electric and magnetic forces exactly cancel when $v = E/B$. (See Examples 27.5 and 27.6.)

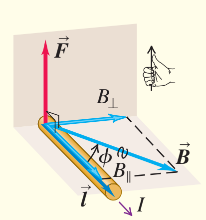
$$R = \frac{mv}{|q|B} \qquad (27.11)$$



Magnetic force on a conductor: A straight segment of a conductor carrying current I in a uniform magnetic field $\vec{B}$ experiences a force $\vec{F}$ that is perpendicular to both $\vec{B}$ and the vector $\vec{l}$, which points in the direction of the current and has magnitude equal to the length of the segment. A similar relationship gives the force $d\vec{F}$ on an infinitesimal current-carrying segment $d\vec{l}$. (See Examples 27.7 and 27.8.)

$$\vec{F} = I\vec{l} \times \vec{B} \qquad (27.19)$$

$$d\vec{F} = I\, d\vec{l} \times \vec{B} \qquad (27.20)$$

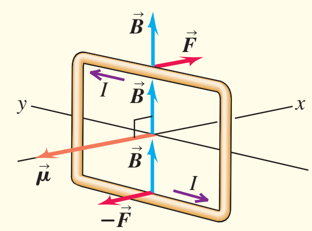


Magnetic torque: A current loop with area A and current I in a uniform magnetic field $\vec{B}$ experiences no net magnetic force, but does experience a magnetic torque of magnitude τ. The vector torque $\vec{\tau}$ can be expressed in terms of the magnetic moment $\vec{\mu} = I\vec{A}$ of the loop, as can the potential energy U of a magnetic moment in a magnetic field $\vec{B}$. The magnetic moment of a loop depends only on the current and the area; it is independent of the shape of the loop. (See Examples 27.9 and 27.10.)

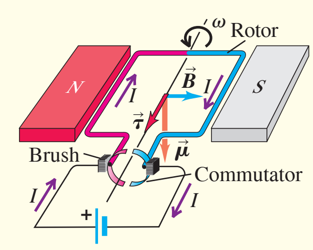
$$\tau = IBA\sin\phi \qquad (27.23)$$

$$\vec{\tau} = \vec{\mu} \times \vec{B} \qquad (27.26)$$

$$U = -\vec{\mu}\cdot\vec{B} = -\mu B\cos\phi \qquad (27.27)$$

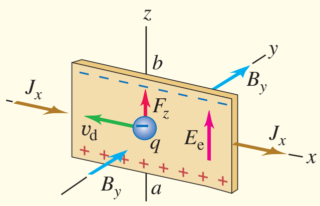


Electric motors: In a dc motor a magnetic field exerts a torque on a current in the rotor. Motion of the rotor through the magnetic field causes an induced emf called a back emf. For a series motor, in which the rotor coil is in series with coils that produce the magnetic field, the terminal voltage is the sum of the back emf and the potential drop Ir across the internal resistance. (See Example 27.11.)



The Hall effect: The Hall effect is a potential difference perpendicular to the direction of current in a conductor, when the conductor is placed in a magnetic field. The Hall potential is determined by the requirement that the associated electric field must just balance the magnetic force on a moving charge. Hall-effect measurements can be used to determine the sign of charge carriers and their concentration n. (See Example 27.12.)

$$nq = \frac{-J_x B_y}{E_z} \tag{27.30}$$



Chapter 27 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLE 27.1 (Section 27.2) before attempting these problems.

VP27.1.1 An electron (charge -1.60×10^{-19} C) moves at 2.20×10^5 m/s through a uniform 1.55 T magnetic field that points in the $+y$-direction. The velocity of the electron lies in the xy-plane and is directed at 40.0° to the $+x$-axis and 50.0° to the $+y$-axis. Find (a) the magnitude and (b) the direction of the magnetic force on the electron.

VP27.1.2 A charged particle moving in the presence of a 0.750 T magnetic field experiences a magnetic force of magnitude 2.50×10^{-10} N. The particle is moving at 1.95×10^5 m/s at an angle of 36.0° to the direction of the magnetic field. Find the magnitude of the charge on the particle.

VP27.1.3 A particle (charge $+3.20 \times 10^{-19}$ C) with velocity $v_x = 3.75 \times 10^5$ m/s is in the presence of a $\vec{B}$ field with only a z-component. The magnetic force on the particle is 2.10×10^{-15} N in the $-y$-direction. Find (a) the magnitude and (b) the direction of $\vec{B}$.

VP27.1.4 Highly ionized iron atoms are found in the sun's outer atmosphere, or corona. One such ion has lost 13 electrons and so has charge $+13e = +2.08 \times 10^{-18}$ C. This ion experiences a magnetic force of 2.30×10^{-15} N when it moves at 7.30×10^3 m/s in the xz-plane, 36.9° from the $+x$-axis and 53.1° from the $+z$-axis. The magnetic field is in the $+z$-direction. Find (a) the magnitude of the magnetic field and (b) the direction of the force on the ion.

Be sure to review EXAMPLES 27.3 and 27.4 (Section 27.4) and EXAMPLES 27.5 and 27.6 (Section 27.5) before attempting these problems.

VP27.6.1 A typical magnetic field in sunspots (highly magnetized regions on the surface of the sun) is 0.300 T. For a proton (charge $+1.60 \times 10^{-19}$ C, mass 1.67×10^{-27} kg) moving at 1.25×10^4 m/s in a direction perpendicular to such a field, find (a) the radius of its circular orbit, (b) its angular speed in its orbit, and (c) the frequency of its orbital motion.

VP27.6.2 In the situation of the previous problem, suppose the 0.300 T magnetic field is in the $+y$-direction and the proton's motion is *not* perpendicular to the field: Initially its velocity has components $v_x = 1.00 \times 10^4$ m/s, $v_y = 7.50 \times 10^3$ m/s, $v_z = 0$. Find (a) the radius of the proton's helical path, (b) how far the proton moves along the helix axis per revolution, and (c) the magnitude of the magnetic force on the proton.

VP27.6.3 A velocity selector uses an electric field $\vec{E} = (2.80 \times 10^4 \text{ N/C})\hat{\imath}$ and a magnetic field $\vec{B} = (0.0350 \text{ T})\hat{k}$. (a) What particle speed will yield zero deflection? (b) In what direction should a charged particle travel through these fields to have zero deflection? (c) A positively charged particle travels through these fields with the speed you found in part (a) and the direction you found in part (b), but the magnetic-field magnitude is now greater than 0.0350 T. In what direction will the particle initially be deflected?

VP27.6.4 You send a beam of oxygen ions through a mass spectrometer like the one shown in Fig. 27.24. The nuclei emerge from the velocity selector at 2.00×10^4 m/s and encounter a magnetic field of magnitude $B' = 0.0500$ T. All the ions have charge $+1.60 \times 10^{-19}$ C, but some are the isotope ^{16}O (mass 2.66×10^{-26} kg) and some are the isotope ^{18}O (mass 2.99×10^{-26} kg). Find the radius of the semicircular path followed by (a) the ^{16}O ions and (b) the ^{18}O ions.

Be sure to review EXAMPLE 27.7 (Section 27.6) before attempting these problems.

VP27.7.1 A straight wire 0.150 m in length carries a current of 3.50 A in the $+x$-direction. The wire is in a uniform 0.0136 T magnetic field in the xy-plane that points in a direction 20.0° from the $-x$-axis and 70.0° from the $+y$-axis. Find (a) the magnitude and (b) the direction of the magnetic force on the wire.

VP27.7.2 A vertical straight wire 25.0 cm in length carries a current. You do not know either the magnitude of the current or whether the current is moving upward or downward. If there is a uniform horizontal magnetic field of 0.0350 T that points due north, the wire experiences a horizontal magnetic force to the west of 0.0140 N. Find (a) the magnitude and (b) the direction of the current.

VP27.7.3 A straight wire carries a current of 1.20 A. The length vector of the wire is $\vec{l} = (0.200 \text{ m})\hat{\imath} + (-0.120 \text{ m})\hat{\jmath}$, and the wire is in a uniform magnetic field $\vec{B} = (0.0175 \text{ T})\hat{k}$. Find (a) the x-component, (b) the y-component, (c) the z-component, and (d) the magnitude of the magnetic force on the wire.

VP27.7.4 A straight wire 0.280 m in length carries a current of 3.40 A. What are the *two* angles between the direction of the current and the direction of a uniform 0.0400 T magnetic field for which the magnetic force on the wire has magnitude 0.0250 N?

BRIDGING PROBLEM Magnetic Torque on a Current-Carrying Ring

A circular ring with area 4.45 cm^2 is carrying a current of 12.5 A. The ring, initially at rest, is immersed in a region of uniform magnetic field given by $\vec{B} = (1.15 \times 10^{-2} \text{ T})(12\hat{\imath} + 3\hat{\jmath} - 4\hat{k})$. The ring is positioned initially such that its magnetic moment is given by $\vec{\mu}_i = \mu(-0.800\hat{\imath} + 0.600\hat{\jmath})$, where μ is the (positive) magnitude of the magnetic moment. (a) Find the initial magnetic torque on the ring. (b) The ring (which is free to rotate around one diameter) is released and turns through an angle of 90.0°, at which point its magnetic moment is given by $\vec{\mu}_f = -\mu\hat{k}$. Determine the decrease in potential energy. (c) If the moment of inertia of the ring about a diameter is $8.50 \times 10^{-7} \text{ kg} \cdot \text{m}^2$, determine the angular speed of the ring as it passes through the second position.

SOLUTION GUIDE

IDENTIFY and SET UP

1. The current-carrying ring acts as a magnetic dipole, so you can use the equations for a magnetic dipole in a uniform magnetic field.
2. There are no nonconservative forces acting on the ring as it rotates, so the sum of its rotational kinetic energy (discussed in Section 9.4) and the potential energy is conserved.

EXECUTE

3. Use the vector expression for the torque on a magnetic dipole to find the answer to part (a). (*Hint:* Review Section 1.10.)
4. Find the change in potential energy from the first orientation of the ring to the second orientation.
5. Use your result from step 4 to find the rotational kinetic energy of the ring when it is in the second orientation.
6. Use your result from step 5 to find the ring's angular speed when it is in the second orientation.

EVALUATE

7. If the ring were free to rotate around *any* diameter, in what direction would the magnetic moment point when the ring is in a state of stable equilibrium?

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q27.1 Can a charged particle move through a magnetic field without experiencing any force? If so, how? If not, why not?

Q27.2 At any point in space, the electric field $\vec{E}$ is defined to be in the direction of the electric force on a positively charged particle at that point. Why don't we similarly define the magnetic field $\vec{B}$ to be in the direction of the magnetic force on a moving, positively charged particle?

Q27.3 Section 27.2 describes a procedure for finding the direction of the magnetic force using your right hand. If you use the same procedure, but with your left hand, will you get the correct direction for the force? Explain.

Q27.4 The magnetic force on a moving charged particle is always perpendicular to the magnetic field $\vec{B}$. Is the trajectory of a moving charged particle always perpendicular to the magnetic field lines? Explain your reasoning.

Q27.5 A charged particle is fired into a cubical region of space where there is a uniform magnetic field. Outside this region, there is no magnetic field. Is it possible that the particle will remain inside the cubical region? Why or why not?

Q27.6 If the magnetic force does no work on a charged particle, how can it have any effect on the particle's motion? Are there other examples of forces that do no work but have a significant effect on a particle's motion?

Q27.7 A charged particle moves through a region of space with constant velocity (magnitude and direction). If the external magnetic field is zero in this region, can you conclude that the external electric field in the region is also zero? Explain. (By "external" we mean fields other than those produced by the charged particle.) If the external electric field is zero in the region, can you conclude that the external magnetic field in the region is also zero?

Q27.8 How might a loop of wire carrying a current be used as a compass? Could such a compass distinguish between north and south? Why or why not?

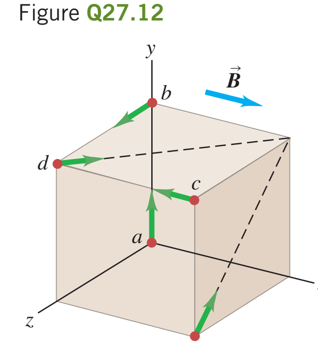
Q27.9 How could the direction of a magnetic field be determined by making only *qualitative* observations of the magnetic force on a straight wire carrying a current?

Q27.10 A loose, floppy loop of wire is carrying current I. The loop of wire is placed on a horizontal table in a uniform magnetic field $\vec{B}$ perpendicular to the plane of the table. This causes the loop of wire to expand into a circular shape while still lying on the table. In a diagram, show all possible orientations of the current I and magnetic field $\vec{B}$ that could cause this to occur. Explain your reasoning.

Q27.11 Several charges enter a uniform magnetic field directed into the page. (a) What path would a positive charge q moving with a velocity of magnitude v follow through the field? (b) What path would a positive charge q moving with a velocity of magnitude $2v$ follow through the field? (c) What path would a negative charge $-q$ moving with a velocity of magnitude v follow through the field? (d) What path would a neutral particle follow through the field?

Q27.12 Each of the lettered points at the corners of the cube in **Fig. Q27.12** represents a positive charge q moving with a velocity of magnitude v in the direction indicated. The region in the figure is in a uniform magnetic field $\vec{B}$, parallel to the x-axis and directed toward the right. Which charges experience a force due to $\vec{B}$? What is the direction of the force on each charge?

Figure Q27.12



Q27.13 A student claims that if lightning strikes a metal flagpole, the force exerted by the earth's magnetic field on the current in the pole can be large enough to bend it. Typical lightning currents are of the order of 10^4 to 10^5 A. Is the student's opinion justified? Explain your reasoning.

Q27.14 Could an accelerator be built in which *all* the forces on the particles, for steering and for increasing speed, are magnetic forces? Why or why not?

Q27.15 The magnetic force acting on a charged particle can never do work because at every instant the force is perpendicular to the velocity. The torque exerted by a magnetic field can do work on a current loop when the loop rotates. Explain how these seemingly contradictory statements can be reconciled.

Q27.16 When the polarity of the voltage applied to a dc motor is reversed, the direction of motion does *not* reverse. Why not? How *could* the direction of motion be reversed?

Q27.17 In a Hall-effect experiment, is it possible that *no* transverse potential difference will be observed? Under what circumstances might this happen?

Q27.18 Hall-effect voltages are much greater for relatively poor conductors (such as germanium) than for good conductors (such as copper), for comparable currents, fields, and dimensions. Why?

EXERCISES

Section 27.2 Magnetic Field

27.1 • A particle with a charge of -1.24×10^{-8} C is moving with instantaneous velocity $\vec{v} = (4.19 \times 10^4 \text{ m/s})\hat{\imath} + (-3.85 \times 10^4 \text{ m/s})\hat{\jmath}$. What is the force exerted on this particle by a magnetic field (a) $\vec{B} = (1.40 \text{ T})\hat{\imath}$ and (b) $\vec{B} = (1.40 \text{ T})\hat{k}$?

27.2 • A particle of mass 0.195 g carries a charge of -2.50×10^{-8} C. The particle is given an initial horizontal velocity that is due north and has magnitude 4.00×10^4 m/s. What are the magnitude and direction of the minimum magnetic field that will keep the particle moving in the earth's gravitational field in the same horizontal, northward direction?

27.3 • In a 1.25 T magnetic field directed vertically upward, a particle having a charge of magnitude 8.50 μC and initially moving northward at 4.75 km/s is deflected toward the east. (a) What is the sign of the charge of this particle? Make a sketch to illustrate how you found your answer. (b) Find the magnetic force on the particle.

27.4 • A particle with mass 1.81×10^{-3} kg and a charge of 1.22×10^{-8} C has, at a given instant, a velocity $\vec{v} = (3.00 \times 10^4 \text{ m/s})\hat{\jmath}$. What are the magnitude and direction of the particle's acceleration produced by a uniform magnetic field $\vec{B} = (1.63 \text{ T})\hat{\imath} + (0.980\ T)\hat{\jmath}$?

27.5 • An electron is moving in the xy-plane. If at time t a magnetic field $B = 0.200$ T in the $+z$-direction exerts a force on the electron equal to $F = 5.50 \times 10^{-18}$ N in the $-y$-direction, what is the velocity (magnitude and direction) of the electron at this instant?

27.6 • An electron moves at 1.40×10^6 m/s through a region in which there is a magnetic field of unspecified direction and magnitude 7.40×10^{-2} T. (a) What are the largest and smallest possible magnitudes of the acceleration of the electron due to the magnetic field? (b) If the actual acceleration of the electron is one-fourth of the largest magnitude in part (a), what is the angle between the electron velocity and the magnetic field?

27.7 •• A group of particles is traveling in a magnetic field of unknown magnitude and direction. You observe that a proton moving at 1.50 km/s in the $+x$-direction experiences a force of 2.25×10^{-16} N in the $+y$-direction, and an electron moving at 4.75 km/s in the $-z$-direction experiences a force of 8.50×10^{-16} N in the $+y$-direction. (a) What are the magnitude and direction of the magnetic field? (b) What are the magnitude and direction of the magnetic force on an electron moving in the $-y$-direction at 3.20 km/s?

27.8 •• **CP** A particle with charge -5.60 nC is moving in a uniform magnetic field $\vec{B} = -(1.25 \text{ T})\hat{k}$. The magnetic force on the particle is measured to be $\vec{F} = -(3.40 \times 10^{-7} \text{ N})\hat{\imath} + (7.40 \times 10^{-7} \text{ N})\hat{\jmath}$. (a) Calculate all the components of the velocity of the particle that you can from this information. (b) Are there components of the velocity that are not determined by the measurement of the force? Explain. (c) Calculate the scalar product $\vec{v} \cdot \vec{F}$. What is the angle between $\vec{v}$ and $\vec{F}$?

Section 27.3 Magnetic Field Lines and Magnetic Flux

27.9 • A circular area with a radius of 6.50 cm lies in the xy-plane. What is the magnitude of the magnetic flux through this circle due to a uniform magnetic field $B = 0.230$ T (a) in the $+z$-direction; (b) at an angle of 53.1° from the $+z$-direction; (c) in the $+y$-direction?

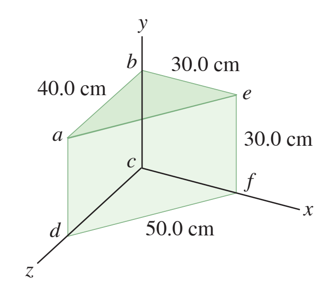
27.10 • A flat, square surface with side length 3.40 cm is in the xy-plane at $z = 0$. Calculate the magnitude of the flux through this surface produced by a magnetic field $\vec{B} = (0.200 \text{ T})\hat{\imath} + (0.300 \text{ T})\hat{\jmath} - (0.500 \text{ T})\hat{k}$.

27.11 •• An open plastic soda bottle with an opening diameter of 2.5 cm is placed on a table. A uniform 1.75 T magnetic field directed upward and oriented 25° from vertical encompasses the bottle. What is the total magnetic flux through the plastic of the soda bottle?

27.12 • A horizontal rectangular surface has dimensions 2.80 cm by 3.20 cm and is in a uniform magnetic field that is directed at an angle of 30.0° above the horizontal. What must the magnitude of the magnetic field be to produce a flux of 3.10×10^{-4} Wb through the surface?

27.13 •• The magnetic field $\vec{B}$ in a certain region is 0.128 T, and its direction is that of the $+z$-axis in **Fig. E27.13**. (a) What is the magnetic flux across the surface $abcd$ in the figure? (b) What is the magnetic flux across the surface $befc$? (c) What is the magnetic flux across the surface $aefd$? (d) What is the net flux through all five surfaces that enclose the shaded volume?

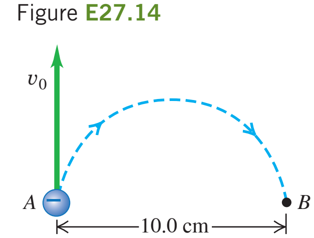
Figure **E27.13**



Section 27.4 Motion of Charged Particles in a Magnetic Field

27.14 •• An electron at point A in **Fig. E27.14** has a speed v_0 of 1.41×10^6 m/s. Find (a) the magnitude and direction of the magnetic field that will cause the electron to follow the semicircular path from A to B, and (b) the time required for the electron to move from A to B.

Figure **E27.14**



27.15 •• Repeat Exercise 27.14 for the case in which the particle is a proton rather than an electron.

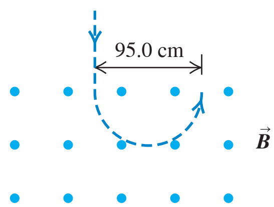
27.16 • An alpha particle (a He nucleus, containing two protons and two neutrons and having a mass of 6.64×10^{-27} kg) traveling horizontally at 35.6 km/s enters a uniform, vertical, 1.80 T magnetic field. (a) What is the diameter of the path followed by this alpha particle? (b) What effect does the magnetic field have on the speed of the particle? (c) What are the magnitude and direction of the acceleration of the alpha particle while it is in the magnetic field? (d) Explain why the speed of the particle does not change even though an unbalanced external force acts on it.

27.17 • **CP** A 150 g ball containing 4.00×10^8 excess electrons is dropped into a 125 m vertical shaft. At the bottom of the shaft, the ball suddenly enters a uniform horizontal magnetic field that has magnitude 0.250 T and direction from east to west. If air resistance is negligibly small, find the magnitude and direction of the force that this magnetic field exerts on the ball just as it enters the field.

27.18 • **BIO** Cyclotrons are widely used in nuclear medicine for producing short-lived radioactive isotopes. These cyclotrons typically accelerate H^- (the *hydride* ion, which has one proton and two electrons) to an energy of 5 MeV to 20 MeV. This ion has a mass very close to that of a proton because the electron mass is negligible—about $\frac{1}{2000}$ of the proton's mass. A typical magnetic field in such cyclotrons is 1.9 T. (a) What is the speed of a 5.0 MeV H^-? (b) If the H^- has energy 5.0 MeV and $B = 1.9$ T, what is the radius of this ion's circular orbit?

27.19 • In an experiment with cosmic rays, a vertical beam of particles that have charge of magnitude $3e$ and mass 12 times the proton mass enters a uniform horizontal magnetic field of 0.250 T and is bent in a semicircle of diameter 95.0 cm, as shown in **Fig. E27.19**. (a) Find the speed of the particles and the sign of their charge. (b) Is it reasonable to ignore the gravity force on the particles? (c) How does the speed of the particles as they enter the field compare to their speed as they exit the field?

Figure **E27.19**

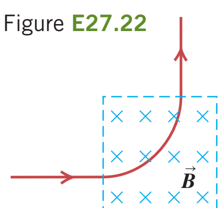


27.20 • A deuteron (the nucleus of an isotope of hydrogen) has a mass of 3.34×10^{-27} kg and a charge of $+e$. The deuteron travels in a circular path with a radius of 6.96 mm in a magnetic field with magnitude 2.50 T. (a) Find the speed of the deuteron. (b) Find the time required for it to make half a revolution. (c) Through what potential difference would the deuteron have to be accelerated to acquire this speed?

27.21 • An electron in the beam of a cathode-ray tube is accelerated by a potential difference of 2.00 kV. Then it passes through a region of transverse magnetic field, where it moves in a circular arc with radius 0.180 m. What is the magnitude of the field?

27.22 •• A beam of protons traveling at 1.20 km/s enters a uniform magnetic field, traveling perpendicular to the field. The beam exits the magnetic field, leaving the field in a direction perpendicular to its original direction (**Fig. E27.22**). The beam travels a distance of 1.18 cm *while in the field*. What is the magnitude of the magnetic field?

Figure **E27.22**



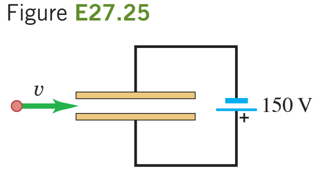
Section 27.5 Applications of Motion of Charged Particles

27.23 • **Crossed $\vec{E}$ and $\vec{B}$ Fields.** A particle with initial velocity $\vec{v}_0 = (5.85 \times 10^3 \text{ m/s})\hat{j}$ enters a region of uniform electric and magnetic fields. The magnetic field in the region is $\vec{B} = -(1.35 \text{ T})\hat{k}$. Calculate the magnitude and direction of the electric field in the region if the particle is to pass through undeflected, for a particle of charge (a) +0.640 nC and (b) −0.320 nC. You can ignore the weight of the particle.

27.24 • (a) What is the speed of a beam of electrons when the simultaneous influence of an electric field of 1.56×10^4 V/m and a magnetic field of 4.62×10^{-3} T, with both fields normal to the beam and to each other, produces no deflection of the electrons? (b) In a diagram, show the relative orientation of the vectors $\vec{v}$, $\vec{E}$, and $\vec{B}$. (c) When the electric field is removed, what is the radius of the electron orbit? What is the period of the orbit?

27.25 •• A 150 V battery is connected across two parallel metal plates of area 28.5 cm^2 and separation 8.20 mm. A beam of alpha particles (charge $+2e$, mass 6.64×10^{-27} kg) is accelerated from rest through a potential difference of 1.75 kV and enters the region between the plates perpendicular to the electric field, as shown in **Fig. E27.25**. What magnitude and direction of magnetic field are needed so that the alpha particles emerge undeflected from between the plates?

Figure **E27.25**



27.26 • In the Bainbridge mass spectrometer (see Fig. 27.22), the magnetic-field magnitude in the velocity selector is 0.510 T, and ions having a speed of 1.82×10^6 m/s pass through undeflected. (a) What is the electric-field magnitude in the velocity selector? (b) If the separation of the plates is 5.20 mm, what is the potential difference between the plates?

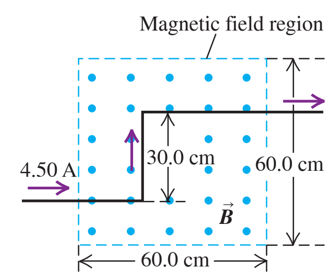
27.27 • Singly ionized (one electron removed) atoms are accelerated and then passed through a velocity selector consisting of perpendicular electric and magnetic fields. The electric field is 155 V/m and the magnetic field is 0.0315 T. The ions next enter a uniform magnetic field of magnitude 0.0175 T that is oriented perpendicular to their velocity. (a) How fast are the ions moving when they emerge from the velocity selector? (b) If the radius of the path of the ions in the second magnetic field is 17.5 cm, what is their mass?

27.28 •• **BIO Ancient Meat Eating.** The amount of meat in prehistoric diets can be determined by measuring the ratio of the isotopes ^{15}N to ^{14}N in bone from human remains. Carnivores concentrate ^{15}N, so this ratio tells archaeologists how much meat was consumed. For a mass spectrometer that has a path radius of 12.5 cm for ^{12}C ions (mass 1.99×10^{-26} kg), find the separation of the ^{14}N (mass 2.32×10^{-26} kg) and ^{15}N (mass 2.49×10^{-26} kg) isotopes at the detector.

Section 27.6 Magnetic Force on a Current-Carrying Conductor

27.29 •• A long wire carrying 4.50 A of current makes two 90° bends, as shown in **Fig. E27.29**. The bent part of the wire passes through a uniform 0.240 T magnetic field directed as shown in the figure and confined to a limited region of space. Find the magnitude and direction of the force that the magnetic field exerts on the wire.

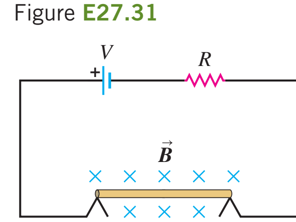
Figure **E27.29**



27.30 • A straight, 2.5 m wire carries a typical household current of 1.5 A (in one direction) at a location where the earth's magnetic field is 0.55 gauss from south to north. Find the magnitude and direction of the force that our planet's magnetic field exerts on this wire if it is oriented so that the current in it is running (a) from west to east, (b) vertically upward, (c) from north to south. (d) Is the magnetic force ever large enough to cause significant effects under normal household conditions?

27.31 • A thin, 50.0-cm-long metal bar with mass 750 g rests on, but is not attached to, two metallic supports in a uniform 0.450 T magnetic field, as shown in **Fig. E27.31**. A battery and a 25.0 Ω resistor in series are connected to the supports. (a) What is the highest voltage the battery can have without breaking the circuit at the supports? (b) The battery voltage has the maximum value calculated in part (a). If the resistor suddenly gets partially short-circuited, decreasing its resistance to 2.0 Ω, find the initial acceleration of the bar.

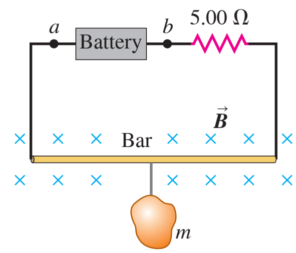
Figure **E27.31**



27.32 •• An electromagnet produces a magnetic field of 0.550 T in a cylindrical region of radius 2.50 cm between its poles. A straight wire carrying a current of 10.8 A passes through the center of this region and is perpendicular to both the axis of the cylindrical region and the magnetic field. What magnitude of force does this field exert on the wire?

27.33 • **Magnetic Balance.** The circuit shown in **Fig. E27.33** is used to make a magnetic balance to weigh objects. The mass m to be measured is hung from the center of the bar that is in a uniform magnetic field of 1.50 T, directed into the plane of the figure. The battery voltage can be adjusted to vary the current in the circuit. The horizontal bar is 60.0 cm long and is made of extremely lightweight material. It is connected to the battery by thin vertical wires that can support no appreciable tension; all the weight of the suspended mass m is supported by the magnetic force on the bar. A resistor with $R = 5.00\ \Omega$ is in series with the bar; the resistance of the rest of the circuit is much less than this. (a) Which point, a or b, should be the positive terminal of the battery? (b) If the maximum terminal voltage of the battery is 175 V, what is the greatest mass m that this instrument can measure?

Figure **E27.33**



27.34 • A straight, vertical wire carries a current of 2.60 A downward in a region between the poles of a large superconducting electromagnet, where the magnetic field has magnitude $B = 0.588$ T and is horizontal. What are the magnitude and direction of the magnetic force on a 1.00 cm section of the wire that is in this uniform magnetic field, if the magnetic field direction is (a) east; (b) south; (c) 30.0° south of west?

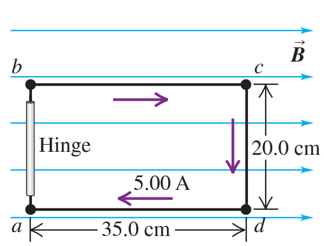
Section 27.7 Force and Torque on a Current Loop

27.35 • A flat circular coil carrying a current of 8.80 A has a magnetic dipole moment of $0.194\ \text{A}\cdot\text{m}^2$ to the left. Its area vector $\vec{A}$ is $4.0\ \text{cm}^2$ to the left. (a) How many turns does the coil have? (b) An observer is on the coil's axis to the left of the coil and is looking toward the coil. Does the observer see a clockwise or counterclockwise current? (c) If a huge 45.0 T external magnetic field directed out of the paper is applied to the coil, what torque (magnitude and direction) results?

27.36 •• The plane of a 5.0 cm × 8.0 cm rectangular loop of wire is parallel to a 0.19 T magnetic field. The loop carries a current of 6.2 A. (a) What torque acts on the loop? (b) What is the magnetic moment of the loop? (c) What is the maximum torque that can be obtained with the same total length of wire carrying the same current in this magnetic field?

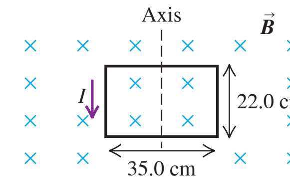
27.37 • The 20.0 cm × 35.0 cm rectangular circuit shown in **Fig. E27.37** is hinged along side ab. It carries a clockwise 5.00 A current and is located in a uniform 1.20 T magnetic field oriented perpendicular to two of its sides, as shown. (a) Draw a clear diagram showing the direction of the force that the magnetic field exerts on each segment of the circuit (ab, bc, etc.). (b) Of the four forces you drew in part (a), decide which ones exert a torque about the hinge ab. Then calculate only those forces that exert this torque. (c) Use your results from part (b) to calculate the torque that the magnetic field exerts on the circuit about the hinge axis ab.

Figure **E27.37**



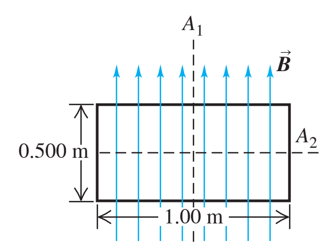
27.38 • A rectangular coil of wire, 22.0 cm by 35.0 cm and carrying a current of 1.95 A, is oriented with the plane of its loop perpendicular to a uniform 1.50 T magnetic field (**Fig. E27.38**). (a) Calculate the net force and torque that the magnetic field exerts on the coil. (b) The coil is rotated through a 30.0° angle about the axis shown, with the left side coming out of the plane of the figure and the right side going into the plane. Calculate the net force and torque that the magnetic field now exerts on the coil. (*Hint:* To visualize this three-dimensional problem, make a careful drawing of the coil as viewed along the rotation axis.)

Figure **E27.38**



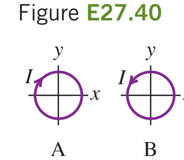
27.39 • **CP** A uniform rectangular coil of total mass 212 g and dimensions 0.500 m × 1.00 m is oriented with its plane parallel to a uniform 3.00 T magnetic field (**Fig. E27.39**). A current of 2.00 A is suddenly started in the coil. (a) About which axis (A_1 or A_2) will the coil begin to rotate? Why? (b) Find the initial angular acceleration of the coil just after the current is started.

Figure **E27.39**



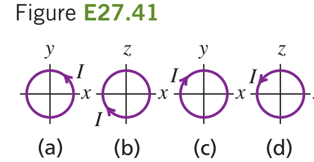
27.40 • Both circular coils A and B (**Fig. E27.40**) have area A and N turns. They are free to rotate about a diameter that coincides with the x-axis. Current I circulates in each coil in the direction shown. There is a uniform magnetic field $\vec{B}$ in the $+z$-direction. (a) What is the direction of the magnetic moment $\vec{\mu}$ for each coil? (b) Explain why the torque on both coils due to the magnetic field is zero, so the coil is in rotational equilibrium. (c) Use Eq. (27.27) to calculate the potential energy for each coil. (d) For each coil, is the equilibrium stable or unstable? Explain.

Figure **E27.40**



27.41 • A circular coil with area A and N turns is free to rotate about a diameter that coincides with the x-axis. Current I is circulating in the coil. There is a uniform magnetic field $\vec{B}$ in the positive y-direction. Calculate the magnitude and direction of the torque $\vec{\tau}$ and the value of the potential energy U, as given in Eq. (27.27), when the coil is oriented as shown in parts (a) through (d) of **Fig. E27.41**.

Figure **E27.41**

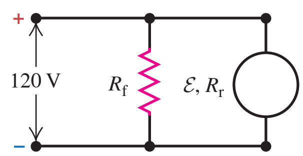


27.42 •• A coil with magnetic moment $1.45\ \text{A}\cdot\text{m}^2$ is oriented initially with its magnetic moment antiparallel to a uniform 0.835 T magnetic field. What is the change in potential energy of the coil when it is rotated 180° so that its magnetic moment is parallel to the field?

Section 27.8 The Direct-Current Motor

27.43 •• In a shunt-wound dc motor with the field coils and rotor connected in parallel (**Fig. E27.43**), the resistance R_f of the field coils is 106 Ω, and the resistance R_r of the rotor is 5.9 Ω. When a potential difference of 120 V is applied to the brushes and the motor is running at full speed delivering mechanical power, the current supplied to it is 4.82 A. (a) What is the current in the field coils? (b) What is the current in the rotor? (c) What is the induced emf developed by the motor? (d) How much mechanical power is developed by this motor?

Figure **E27.43**

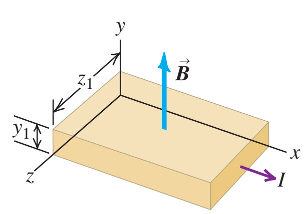


27.44 • A dc motor with its rotor and field coils connected in series has an internal resistance of 3.2 Ω. When the motor is running at full load on a 120 V line, the emf in the rotor is 105 V. (a) What is the current drawn by the motor from the line? (b) What is the power delivered to the motor? (c) What is the mechanical power developed by the motor?

Section 27.9 The Hall Effect

27.45 • **Figure E27.45** shows a portion of a silver ribbon with $z_1 = 11.8$ mm and $y_1 = 0.23$ mm, carrying a current of 120 A in the $+x$-direction. The ribbon lies in a uniform magnetic field, in the y-direction, with magnitude 0.95 T. Apply the simplified model of the Hall effect presented in Section 27.9. If there are 5.85×10^{28} free electrons per cubic meter, find (a) the magnitude of the drift velocity of the electrons in the x-direction; (b) the magnitude and direction of the electric field in the z-direction due to the Hall effect; (c) the Hall emf.

Figure **E27.45**



27.46 • Let Fig. E27.45 represent a strip of an unknown metal of the same dimensions as those of the silver ribbon in Exercise 27.45. When the magnetic field is 2.29 T and the current is 78.0 A, the Hall emf is found to be 131 μV. What does the simplified model of the Hall effect presented in Section 27.9 give for the density of free electrons in the unknown metal?

PROBLEMS

27.47 •• **CP** A small particle with positive charge $q = +3.75 \times 10^{-4}$ C and mass $m = 5.00 \times 10^{-5}$ kg is moving in a region of uniform electric and magnetic fields. The magnetic field is $B = 4.00$ T in the $+z$-direction. The electric field is also in the $+z$-direction and has magnitude $E = 60.0$ N/C. At time $t = 0$ the particle is on the y-axis at $y = +1.00$ m and has velocity $v = 30.0$ m/s in the $+x$-direction. Neglect gravity. (a) What are the x-, y-, and z-coordinates of the particle at $t = 0.0200$ s? (b) What is the speed of the particle at $t = 0.0200$ s?

27.48 • A particle with charge 7.26×10^{-8} C is moving in a region where there is a uniform 0.650 T magnetic field in the $+x$-direction. At a particular instant, the velocity of the particle has components $v_x = -1.68 \times 10^4$ m/s, $v_y = -3.11 \times 10^4$ m/s, and $v_z = 5.85 \times 10^4$ m/s. What are the components of the force on the particle at this time?

27.49 ••• **CP** **Fusion Reactor.** If two deuterium nuclei (charge $+e$, mass 3.34×10^{-27} kg) get close enough together, the attraction of the strong nuclear force will fuse them to make an isotope of helium, releasing vast amounts of energy. The range of this force is about 10^{-15} m. This is the principle behind the fusion reactor. The deuterium nuclei are moving much too fast to be contained by physical walls, so they are confined magnetically. (a) How fast would two nuclei have to move so that in a head-on collision they would get close enough to fuse? (Assume their speeds are equal. Treat the nuclei as point charges, and assume that a separation of 1.0×10^{-15} m is required for fusion.) (b) What strength magnetic field is needed to make deuterium nuclei with this speed travel in a circle of diameter 2.50 m?

27.50 •• **Magnetic Moment of the Hydrogen Atom.** In the Bohr model of the hydrogen atom (see Section 39.3), in the lowest energy state the electron orbits the proton at a speed of 2.2×10^6 m/s in a circular orbit of radius 5.3×10^{-11} m. (a) What is the orbital period of the electron? (b) If the orbiting electron is considered to be a current loop, what is the current I? (c) What is the magnetic moment of the atom due to the motion of the electron?

27.51 •• You measure the charge-to-mass ratio q/m for a particle with positive charge in the following way: The particle starts from rest, is accelerated through a potential difference ΔV, and attains a velocity with magnitude v. It then enters a region of uniform magnetic field $B = 0.200$ T that is directed perpendicular to the velocity; the particle moves in a path that is an arc of a circle of radius R. You measure R as a function of ΔV. You plot your data as R^2 (in units of m^2) versus ΔV (in V) and find that the values lie close to a straight line that has slope 1.04×10^{-6} m^2/V. What is the value of q/m for this particle?

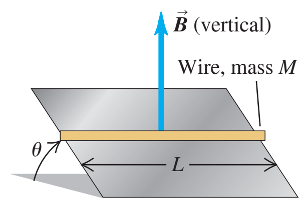
27.52 • The magnetic poles of a small cyclotron produce a magnetic field with magnitude 0.85 T. The poles have a radius of 0.40 m, which is the maximum radius of the orbits of the accelerated particles. (a) What is the maximum energy to which protons ($q = 1.60 \times 10^{-19}$ C, $m = 1.67 \times 10^{-27}$ kg) can be accelerated by this cyclotron? Give your answer in electron volts and in joules. (b) What is the time for one revolution of a proton orbiting at this maximum radius? (c) What would the magnetic-field magnitude have to be for the maximum energy to which a proton can be accelerated to be twice that calculated in part (a)? (d) For $B = 0.85$ T, what is the maximum energy to which alpha particles ($q = 3.20 \times 10^{-19}$ C, $m = 6.64 \times 10^{-27}$ kg) can be accelerated by this cyclotron? How does this compare to the maximum energy for protons?

27.53 •• Suppose the electric field between the plates in Fig. 27.22 is 1.88×10^4 V/m and the magnetic field in both regions is 0.682 T. If the source contains the three isotopes of krypton, ^{82}Kr, ^{84}Kr, and ^{86}Kr, and the ions are singly charged, find the distance between the lines formed by the three isotopes on the particle detector. Assume the atomic masses of the isotopes (in atomic mass units) are equal to their mass numbers, 82, 84, and 86. (One atomic mass unit $= 1\ \text{u} = 1.66 \times 10^{-27}$ kg.)

27.54 •• **Mass Spectrograph.** A mass spectrograph is used to measure the masses of ions, or to separate ions of different masses (see Section 27.5). In one design for such an instrument, ions with mass m and charge q are accelerated through a potential difference V. They then enter a uniform magnetic field that is perpendicular to their velocity, and they are deflected in a semicircular path of radius R. A detector measures where the ions complete the semicircle and from this it is easy to calculate R. (a) Derive the equation for calculating the mass of the ion from measurements of B, V, R, and q. (b) What potential difference V is needed so that singly ionized ^{12}C atoms will have $R = 50.0$ cm in a 0.150 T magnetic field? (c) Suppose the beam consists of a mixture of ^{12}C and ^{14}C ions. If v and B have the same values as in part (b), calculate the separation of these two isotopes at the detector. Do you think that this beam separation is sufficient for the two ions to be distinguished? (Make the assumption described in Problem 27.53 for the masses of the ions.)

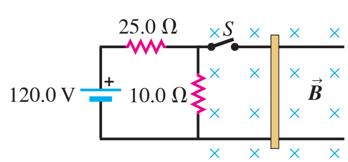
27.55 •• A straight piece of conducting wire with mass M and length L is placed on a frictionless incline tilted at an angle θ from the horizontal (**Fig. P27.55**). There is a uniform, vertical magnetic field $\vec{B}$ at all points (produced by an arrangement of magnets not shown in the figure). To keep the wire from sliding down the incline, a voltage source is attached to the ends of the wire. When just the right amount of current flows through the wire, the wire remains at rest. Determine the magnitude and direction of the current in the wire that will cause the wire to remain at rest. Copy the figure and draw the direction of the current on your copy. In addition, show in a free-body diagram all the forces that act on the wire.

Figure **P27.55**



27.56 •• **CP** A 2.60 N metal bar, 0.850 m long and having a resistance of 10.0 Ω, rests horizontally on conducting wires connecting it to the circuit shown in **Fig. P27.56**. The bar is in a uniform, horizontal, 1.60 T magnetic field and is not attached to the wires in the circuit. What is the acceleration of the bar just after the switch S is closed?

Figure **P27.56**

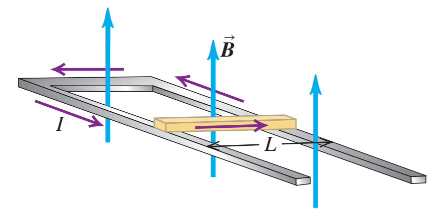


27.57 •• **BIO** **Determining Diet.** One method for determining the amount of corn in early Native American diets is the *stable isotope ratio analysis* (SIRA) technique. As corn photosynthesizes, it concentrates the isotope carbon-13, whereas most other plants concentrate carbon-12. Overreliance on corn consumption can then be correlated with certain diseases, because corn lacks the essential amino acid lysine. Archaeologists use a mass spectrometer to separate the ^{12}C and ^{13}C isotopes in samples of human remains. Suppose you use a velocity selector to obtain singly ionized (missing one electron) atoms of speed 8.50 km/s, and you want to bend them within a uniform magnetic field in a semicircle of diameter 25.0 cm for the ^{12}C. The measured masses of these isotopes are 1.99×10^{-26} kg (^{12}C) and 2.16×10^{-26} kg (^{13}C). (a) What strength of magnetic field is required? (b) What is the diameter of the ^{13}C semicircle? (c) What is the separation of the ^{12}C and ^{13}C ions at the detector at the end of the semicircle? Is this distance large enough to be easily observed?

27.58 •• **CP** A plastic circular loop has radius R, and a positive charge q is distributed uniformly around the circumference of the loop. The loop is then rotated around its central axis, perpendicular to the plane of the loop, with angular speed ω. If the loop is in a region where there is a uniform magnetic field $\vec{B}$ directed parallel to the plane of the loop, calculate the magnitude of the magnetic torque on the loop.

27.59 •• **CP** **An Electromagnetic Rail Gun.** A conducting bar with mass m and length L slides over horizontal rails that are connected to a voltage source. The voltage source maintains a constant current I in the rails and bar, and a constant, uniform, vertical magnetic field $\vec{B}$ fills the region between the rails (**Fig. P27.59**). (a) Find the magnitude and direction of the net force on the conducting bar. Ignore friction, air resistance, and electrical resistance. (b) If the bar has mass m, find the distance d that the bar must move along the rails from rest to attain speed v. (c) It has been suggested that rail guns based on this principle could accelerate payloads into earth orbit or beyond. Find the distance the bar must travel along the rails if it is to reach the escape speed for the earth (11.2 km/s). Let $B = 0.80$ T, $I = 2.0 \times 10^3$ A, $m = 25$ kg, and $L = 50$ cm. For simplicity assume the net force on the object is equal to the magnetic force, as in parts (a) and (b), even though gravity plays an important role in an actual launch in space.

Figure **P27.59**



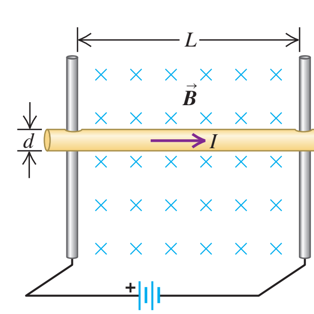
27.60 • A wire 25.0 cm long lies along the z-axis and carries a current of 7.40 A in the $+z$-direction. The magnetic field is uniform and has components $B_x = -0.242$ T, $B_y = -0.985$ T, and $B_z = -0.336$ T. (a) Find the components of the magnetic force on the wire. (b) What is the magnitude of the net magnetic force on the wire?

27.61 •• **CALC** Point a is on the $+y$-axis at $y = +0.200$ m and point b is on the $+x$-axis at $x = +0.200$ m. A wire in the shape of a circular arc of radius 0.200 m and centered on the origin goes from a to b and carries current $I = 5.00$ A in the direction from a to b. (a) If the wire is in a uniform magnetic field $B = 0.800$ T in the $+z$-direction, what are the magnitude and direction of the net force that the magnetic field exerts on the wire segment? (b) What are the magnitude and direction of the net force on the wire if the field is $B = 0.800$ T in the $+x$-direction?

27.62 •• Ceiling fans use electric motors that involve stationary permanent magnets (called stators) attached to the central hub and with a typical strength of 1.0 T, that supply torque to current-carrying coils (called rotors) fixed to the fan blades. Think of a ceiling fan at your home or workplace. (a) Estimate the diameter of the central hub at the inner edge of the blades. (b) If 12 rotors are fixed around the central hub, and if the diameters of the rotors make up one-third of the circumference of the hub, what is the diameter of a single rotor? (c) A typical ceiling fan supplies 1.4 N · m of torque. If this torque is supplied equally by the 12 rotors in the presence of 1.0 T stator fields parallel to the axis of each rotor, estimate the magnitude of the magnetic moment of each rotor. (d) The rotors are connected in parallel, so each receives $\frac{1}{12}$ of the fan current, which is typically 0.75 A. Estimate the number of windings in each rotor.

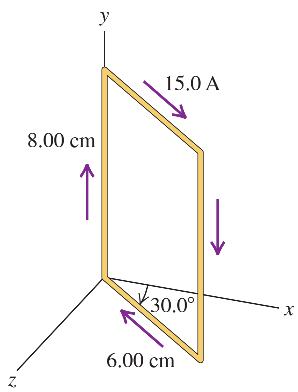
27.63 •• Earth's magnetic field near the ground is typically 0.50 G (0.50 gauss), where 1 G = 10^{-4} tesla. In temperate northern latitudes, this field is inclined downward at an angle of approximately 45°. Consider the feasibility of using the earth's magnetic field to enable the levitation device shown in **Fig. P27.63** (next page). (a) Estimate the current needed to lift a copper bar with diameter $d = 5.0$ mm using the horizontal component of the earth's magnetic field. The density of copper is 8900 kg/m^3. (b) A copper bar with diameter 5.0 mm will melt if it carries a current greater than the fusing current of 900 A. Is it feasible for our device? (c) Estimate the minimum strength of the magnetic field needed to levitate our copper bar. (d) Suppose we use easily obtainable 1.0 T permanent magnets to supply the horizontal magnetic field and suppose our bar has length 10 cm. Estimate how much extra weight we could levitate with our device using a current of 10 A.

Figure **P27.63**



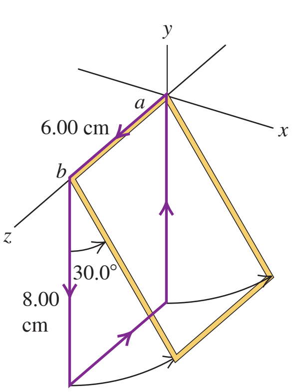
27.64 •• The rectangular loop shown in **Fig. P27.64** is pivoted about the y-axis and carries a current of 15.0 A in the direction indicated. (a) If the loop is in a uniform magnetic field with magnitude 0.48 T in the $+x$-direction, find the magnitude and direction of the torque required to hold the loop in the position shown. (b) Repeat part (a) for the case in which the field is in the $-z$-direction. (c) For each of the above magnetic fields, what torque would be required if the loop were pivoted about an axis through its center, parallel to the y-axis?

Figure **P27.64**



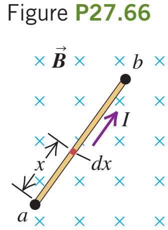
27.65 •• **CP** The rectangular loop of wire shown in **Fig. P27.65** has a mass of 0.15 g per centimeter of length and is pivoted about side ab on a frictionless axis. The current in the wire is 8.2 A in the direction shown. Find the magnitude and direction of the magnetic field parallel to the y-axis that will cause the loop to swing up until its plane makes an angle of 30.0° with the yz-plane.

Figure **P27.65**



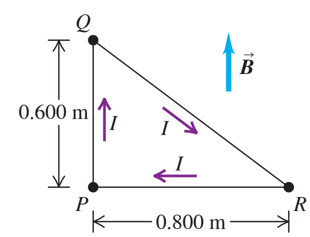
27.66 •• **CALC** A uniform bar of length L carries a current I in the direction from point a to point b (**Fig. P27.66**). The bar is in a uniform magnetic field that is directed into the page. Consider the torque about an axis perpendicular to the bar at point a that is due to the force that the magnetic field exerts on the bar. (a) Suppose that an infinitesimal section of the bar has length dx and is located a distance x from point a. Calculate the torque $d\tau$ about point a due to the magnetic force on this infinitesimal section. (b) Use $\tau = \int_a^b d\tau$ to calculate the total torque τ on the bar. (c) Show that τ is the same as though all of the magnetic force acted at the midpoint of the bar.

Figure **P27.66**



27.67 •• The loop of wire shown in **Fig. P27.67** forms a right triangle and carries a current $I = 5.00$ A in the direction shown. The loop is in a uniform magnetic field that has magnitude $B = 3.00$ T and the same direction as the current in side PQ of the loop. (a) Find the force exerted by the magnetic field on each side of the triangle. If the force is not zero, specify its direction. (b) What is the net force on the loop? (c) The loop is pivoted about an axis that lies along side PR. Use the forces calculated in part (a) to calculate the torque on each side of the loop (see Problem 27.66). (d) What is the magnitude of the net torque on the loop? Calculate the net torque from the torques calculated in part (c) and also from Eq. (27.28). Do these two results agree? (e) Is the net torque directed to rotate point Q into the plane of the figure or out of the plane of the figure?

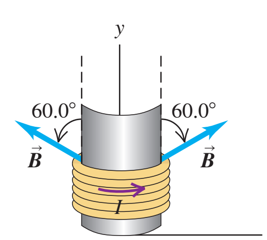
Figure **P27.67**



27.68 •• Repeat Problem 27.67 but with the magnetic field $B = 3.00$ T directed into the page in Fig. P27.67.

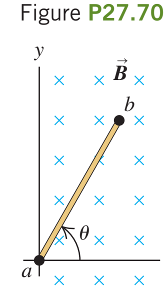
27.69 •• **CALC** **A Voice Coil.** It was shown in Section 27.7 that the net force on a current loop in a *uniform* magnetic field is zero. The magnetic force on the voice coil of a loudspeaker (see Fig. 27.28) is nonzero because the magnetic field at the coil is not uniform. A voice coil in a loudspeaker has 50 turns of wire and a diameter of 1.56 cm, and the current in the coil is 0.950 A. Assume that the magnetic field at each point of the coil has a constant magnitude of 0.220 T and is directed at an angle of 60.0° outward from the normal to the plane of the coil (**Fig. P27.69**). Let the axis of the coil be in the y-direction. The current in the coil is in the direction shown (counterclockwise as viewed from a point above the coil on the y-axis). Calculate the magnitude and direction of the net magnetic force on the coil.

Figure **P27.69**



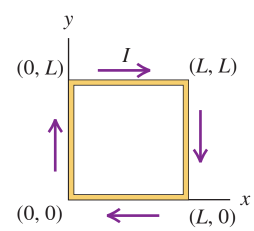
27.70 •• **CP** A uniform bar has mass 0.0120 kg and is 30.0 cm long. It pivots without friction about an axis perpendicular to the bar at point a (**Fig. P27.70**). The gravitational force on the bar acts in the $-y$-direction. The bar is in a uniform magnetic field that is directed into the page and has magnitude $B = 0.150$ T. (a) What must be the current I in the bar for the bar to be in rotational equilibrium when it is at an angle $\theta = 30.0°$ above the horizontal? Use your result from Problem 27.66. (b) For the bar to be in rotational equilibrium, should I be in the direction from a to b or b to a?

Figure **P27.70**



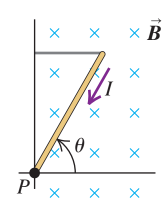
27.71 •• **CALC** **Force on a Current Loop in a Nonuniform Magnetic Field.** It was shown in Section 27.7 that the net force on a current loop in a *uniform* magnetic field is zero. But what if $\vec{B}$ is *not* uniform? **Figure P27.71** shows a square loop of wire that lies in the xy-plane. The loop has corners at $(0, 0)$, $(0, L)$, $(L, 0)$, and (L, L) and carries a constant current I in the clockwise direction. The magnetic field has no x-component but has both y- and z-components: $\vec{B} = (B_0 z/L)\hat{\jmath} + (B_0 y/L)\hat{k}$, where B_0 is a positive constant. (a) Sketch the magnetic field lines in the yz-plane. (b) Find the magnitude and direction of the magnetic force exerted on each of the sides of the loop by integrating Eq. (27.20). (c) Find the magnitude and direction of the net magnetic force on the loop.

Figure P27.71



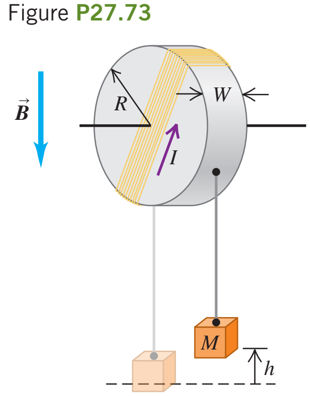
27.72 •• **CP** The lower end of the thin uniform rod in **Fig. P27.72** is attached to the floor by a frictionless hinge at point P. The rod has mass 0.0840 kg and length 18.0 cm and is in a uniform magnetic field $B = 0.120$ T that is directed into the page. The rod is held at an angle $\theta = 53.0°$ above the horizontal by a horizontal string that connects the top of the rod to the wall. The rod carries a current $I = 12.0$ A in the direction toward P. Calculate the tension in the string. Use your result from Problem 27.66 to calculate the torque due to the magnetic-field force.

Figure P27.72



27.73 ••• **CP** A magnetic lift uses a cylinder with radius R and width W wrapped by conducting wire N times on a diameter, as shown in **Fig. P27.73**. A uniform external magnetic field $\vec{B}$ points downward while current I flows in the sense indicated. A mass M is hung from the cylinder by a cable attached to its rim *on the axis of the coil*. The height above the lowest possible position of the mass is h. (a) What is the minimum current I for which the mass can be suspended with $h > 0$? (b) What is the height of the mass $h = h_{top}$ when the cable attachment rotates one-half turn to the top of the cylinder, in terms of R? Define a dimensionless parameter $\sigma = 2NIWB/Mg$ and express subsequent results in terms of σ and other given quantities. (c) What is the net torque on the cylinder if $0 \le h \le R$? (d) What is the net torque on the cylinder if $R \le h \le h_{top}$? (e) Define potential energy $U(h)$ that is zero when $h = 0$. What is the potential energy for $0 \le h \le R$? (f) What is the potential energy for $R \le h \le h_{top}$? (g) If $\sigma > 1$, for what value of h will the mass remain suspended motionless? (h) For what values of σ will the cylinder rotate more than 180° if the mass is released from rest at $h = h_{top}$? (*Hint:* This corresponds to $U(h_{top}) > 0$.)

Figure P27.73



27.74 ••• **CP CALC** A particle with charge q and mass m is dropped at time $t = 0$ from rest at its origin in a region of constant magnetic field $\vec{B}$ that points horizontally. What happens? To answer, construct a Cartesian coordinate system with the y-axis pointing downward and the z-axis pointing in the direction of the magnetic field. At time $t \ge 0$ the particle has velocity $\vec{v} = v_x\hat{\imath} + v_y\hat{\jmath}$. The net force $\vec{F} = F_x\hat{\imath} + F_y\hat{\jmath}$ on the particle is the vector sum of its weight and the magnetic force. (a) Using Newton's second law, write equations for a_x and a_y, where $\vec{a} = a_x\hat{\imath} + a_y\hat{\jmath}$ is the acceleration of the particle. (b) Differentiate the second of these equations with respect to time. Then substitute your expression for $a_x = dv_x/dt$ to determine an equation for $dv_y{}^2/dt^2$ in terms of v_y. (c) This result shows that v_y is a simple harmonic oscillator. Use the initial conditions to determine $v_y(t)$. Write your answer in terms of the angular frequency $\omega = qB/m$. Note that $(dv_y/dt)_0 = g$. (d) Substitute your result for $v_y(t)$ into your equation for dv_x/dt. Integrate using the initial conditions to determine $v_x(t)$. (e) Integrate your expressions for $v_x(t)$ and $v_y(t)$ to determine $x(t)$ and $y(t)$. (f) If $m = 1.00$ mg, $q = 19.6\ \mu$C, and $B = 10.0$ T, what maximum vertical distance does the particle drop before returning upward?

27.75 •• A circular loop of wire with area A lies in the xy-plane. As viewed along the z-axis looking in the $-z$-direction toward the origin, a current I is circulating clockwise around the loop. The torque produced by an external magnetic field $\vec{B}$ is given by $\vec{\tau} = D(4\hat{\imath} - 3\hat{\jmath})$, where D is a positive constant, and for this orientation of the loop the magnetic potential energy $U = -\vec{\mu}\cdot\vec{B}$ is negative. The magnitude of the magnetic field is $B_0 = 13D/IA$. (a) Determine the vector magnetic moment of the current loop. (b) Determine the components B_x, B_y, and B_z of $\vec{B}$.

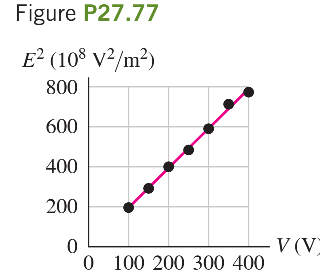
27.76 •• **DATA** You are using a type of mass spectrometer to measure charge-to-mass ratios of atomic ions. In the device, atoms are ionized with a beam of electrons to produce positive ions, which are then accelerated through a potential difference V. (The final speed of the ions is great enough that you can ignore their initial speed.) The ions then enter a region in which a uniform magnetic field $\vec{B}$ is perpendicular to the velocity of the ions and has magnitude $B = 0.250$ T. In this $\vec{B}$ region, the ions move in a semicircular path of radius R. You measure R as a function of the accelerating voltage V for one particular atomic ion:

V (kV)	10.0	12.0	14.0	16.0	18.0
R (cm)	19.9	21.8	23.6	25.2	26.8

(a) How can you plot the data points so that they will fall close to a straight line? Explain. (b) Construct the graph described in part (a). Use the slope of the best-fit straight line to calculate the charge-to-mass ratio (q/m) for the ion. (c) For $V = 20.0$ kV, what is the speed of the ions as they enter the $\vec{B}$ region? (d) If ions that have $R = 21.2$ cm for $V = 12.0$ kV are singly ionized, what is R when $V = 12.0$ kV for ions that are doubly ionized?

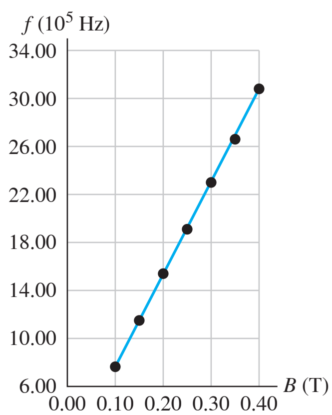
27.77 •• **DATA** You are a research scientist working on a high-energy particle accelerator. Using a modern version of the Thomson e/m apparatus, you want to measure the mass of a muon (a fundamental particle that has the same charge as an electron but greater mass). The magnetic field between the two charged plates is 0.340 T. You measure the electric field for zero particle deflection as a function of the accelerating potential V. This potential is large enough that you can assume the initial speed of the muons to be zero. **Figure P27.77** is an E^2-versus-V graph of your data. (a) Explain why the data points fall close to a straight line. (b) Use the graph in Fig. P27.77 to calculate the mass m of a muon. (c) If the two charged plates are separated by 6.00 mm, what must be the voltage between the plates in order for the electric field between the plates to be 2.00×10^5 V/m? Assume that the dimensions of the plates are much larger than the separation between them. (d) When $V = 400$ V, what is the speed of the muons as they enter the region between the plates?

Figure P27.77



27.78 •• **DATA** You are a technician testing the operation of a cyclotron. An alpha particle in the device moves in a circular path in a magnetic field $\vec{B}$ that is directed perpendicular to the path of the alpha particle. You measure the number of revolutions per second (the frequency f) of the alpha particle as a function of the magnetic field strength B. **Figure P27.78** shows your results and the best straight-line fit to your data. (a) Use the graph in Fig. P27.78 to calculate the charge-to-mass ratio of the alpha particle, which has charge $+2e$. On the basis of your data, what is the mass of an alpha particle? (b) With $B = 0.300$ T, what are the cyclotron frequencies f of a proton and of an electron? How do these f values compare to the frequency of an alpha particle? (c) With $B = 0.300$ T, what speed and kinetic energy does an alpha particle have if the radius of its path is 12.0 cm?

Figure **P27.78**



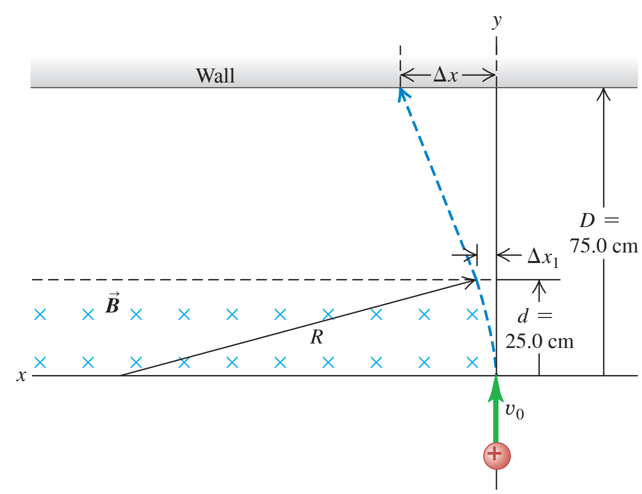
CHALLENGE PROBLEMS

27.79 ••• **CALC** Determine the magnetic moment $\vec{\mu}$ of a spherical shell with radius R and uniform charge Q rotating with angular speed $\vec{\Omega} = \omega\hat{k}$. Use the following steps: (a) Consider a coordinate system with the origin at the center of the sphere. Parameterize each latitude on the sphere with the angle θ measured from the positive z-axis. There is a circular current loop at each value of θ for $0 \leq \theta \leq \pi$. What is the radius of the loop at latitude θ? (b) The differential current carried by that loop is $dI = \sigma v\, dW$, where σ is the charge density of the sphere, v is the tangential speed of the loop, and $dW = R\, d\theta$ is its differential width. Express dI in terms of σ, R, ω, θ, and $d\theta$. (c) The differential magnetic moment of the loop is $d\mu = A\, dI$, where A is the area enclosed by the loop. Express $d\mu$ in terms of R, ω, θ, and $d\theta$. (d) Integrate over the sphere to determine the magnetic moment. Express your result as a vector; use the total charge Q rather than the charge density σ. (e) If the sphere is in a uniform magnetic field $\vec{B} = (\sin\alpha\hat{\imath} + \cos\alpha\hat{\jmath})B$, what is the torque on the sphere?

27.80 ••• An electron traveling from the sun as part of the solar wind strikes the earth's magnetosphere at latitude 80.0° N in a region where the magnetic field has a strength of 15.0 μT and is directed toward the earth's center. The electron has a speed of 400 km/s and is directed toward the earth's axis parallel to the equator. Magnetic forces send the electron on a helical trajectory. (a) What is the radius of this helix? (b) With what speed does the electron approach the surface of the earth? (c) If you are looking downward toward the earth from space, is the electron's motion clockwise or counterclockwise? (d) What is the frequency of the motion? (e) This electron strikes the ionosphere, where it is further accelerated by an electric field with strength 20.0 mV/m directed northward parallel to the earth's surface. What is the electron's new speed after it has been deflected 100 km southward by this field? (f) By what factor has its kinetic energy been increased by the electric field?

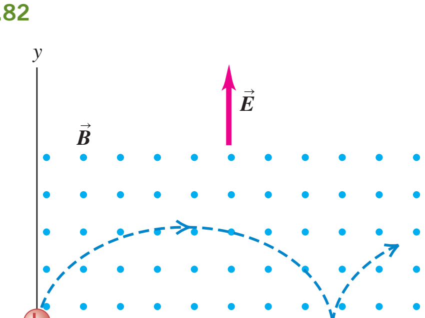
27.81 ••• A particle with charge 2.15 μC and mass 3.20×10^{-11} kg is initially traveling in the $+y$-direction with a speed $v_0 = 1.45 \times 10^5$ m/s. It then enters a region containing a uniform magnetic field that is directed into, and perpendicular to, the page in **Fig. P27.81**. The magnitude of the field is 0.420 T. The region extends a distance of 25.0 cm along the initial direction of travel; 75.0 cm from the point of entry into the magnetic field region is a wall. The length of the field-free region is thus 50.0 cm. When the charged particle enters the magnetic field, it follows a curved path whose radius of curvature is R. It then leaves the magnetic field after a time t_1, having been deflected a distance Δx_1. The particle then travels in the field-free region and strikes the wall after undergoing a total deflection Δx. (a) Determine the radius R of the curved part of the path. (b) Determine t_1, the time the particle spends in the magnetic field. (c) Determine Δx_1, the horizontal deflection at the point of exit from the field. (d) Determine Δx, the total horizontal deflection.

Figure **P27.81**



27.82 ••• **CP** **A Cycloidal Path.** A particle with mass m and positive charge q starts from rest at the origin shown in **Fig. P27.82**. There is a uniform electric field $\vec{E}$ in the $+y$-direction and a uniform magnetic field $\vec{B}$ directed out of the page. It is shown in more advanced books that the path is a *cycloid* whose radius of curvature at the top points is twice the y-coordinate at that level. (a) Explain why the path has this general shape and why it is repetitive. (b) Prove that the speed at any point is equal to $\sqrt{2qEy/m}$. (*Hint:* Use energy conservation.) (c) Applying Newton's second law at the top point and taking as given that the radius of curvature here equals $2y$, prove that the speed at this point is $2E/B$.

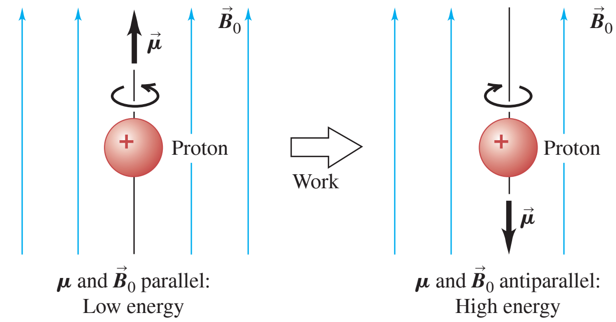
Figure **P27.82**



MCAT-STYLE PASSAGE PROBLEMS

BIO Magnetic Fields and MRI. *Magnetic resonance imaging* (MRI) is a powerful imaging method that, unlike x-ray imaging, allows sharp images of soft tissue to be made without exposing the patient to potentially damaging radiation. A rudimentary understanding of this method can be achieved by the relatively simple application of the classical (that is, non-quantum) physics of magnetism. The starting point for MRI is *nuclear magnetic resonance* (NMR), a technique that depends on the fact that protons in the atomic nucleus have a magnetic field $\vec{B}$. The origin of the proton's magnetic field is the spin of the proton. Being charged, the spinning proton constitutes an electric current analogous to a wire loop through which current flows. Like the wire loop, the proton has a magnetic moment $\vec{\mu}$; thus it will experience a torque when it is subjected to an external magnetic field $\vec{B}_0$. The magnitude of $\vec{\mu}$ is about 1.4×10^{-26} J/T. The proton can be thought of as being in one of two states, with $\vec{\mu}$ oriented parallel or antiparallel to the applied magnetic field, and work must be done to flip the proton from the low-energy state to the high-energy state, as the accompanying figure below shows.

An important consideration is that the net magnetic field of any nucleus, except for that of hydrogen (which has a proton only), consists of contributions from both protons and neutrons. If a nucleus has an even number of protons and neutrons, they will pair in such a way that half of the protons have spins in one orientation and half have spins in the other orientation. Thus the net magnetic moment of the nucleus is zero. Only nuclei with a net magnetic moment are candidates for MRI. Hydrogen is the atom that is most commonly imaged.



27.83 If a proton is exposed to an external magnetic field of 2 T that has a direction perpendicular to the axis of the proton's spin, what will be the torque on the proton? (a) 0; (b) 1.4×10^{-26} N·m; (c) 2.8×10^{-26} N·m; (d) 0.7×10^{-26} N·m.

27.84 Which of following elements is a candidate for MRI? (a) $^{12}C_6$; (b) $^{16}O_8$; (c) $^{40}Ca_{20}$; (d) $^{31}P_{15}$.

27.85 The large magnetic fields used in MRI can produce forces on electric currents within the human body. This effect has been proposed as a possible method for imaging "biocurrents" flowing in the body, such as the current that flows in individual nerves. For a magnetic field strength of 2 T, estimate the magnitude of the maximum force on a 1-mm-long segment of a single cylindrical nerve that has a diameter of 1.5 mm. Assume that the entire nerve carries a current due to an applied voltage of 100 mV (that of a typical action potential). The resistivity of the nerve is 0.6 Ω·m. (a) 6×10^{-7} N; (b) 1×10^{-6} N; (c) 3×10^{-4} N; (d) 0.3 N.

ANSWERS

Chapter Opening Question ?

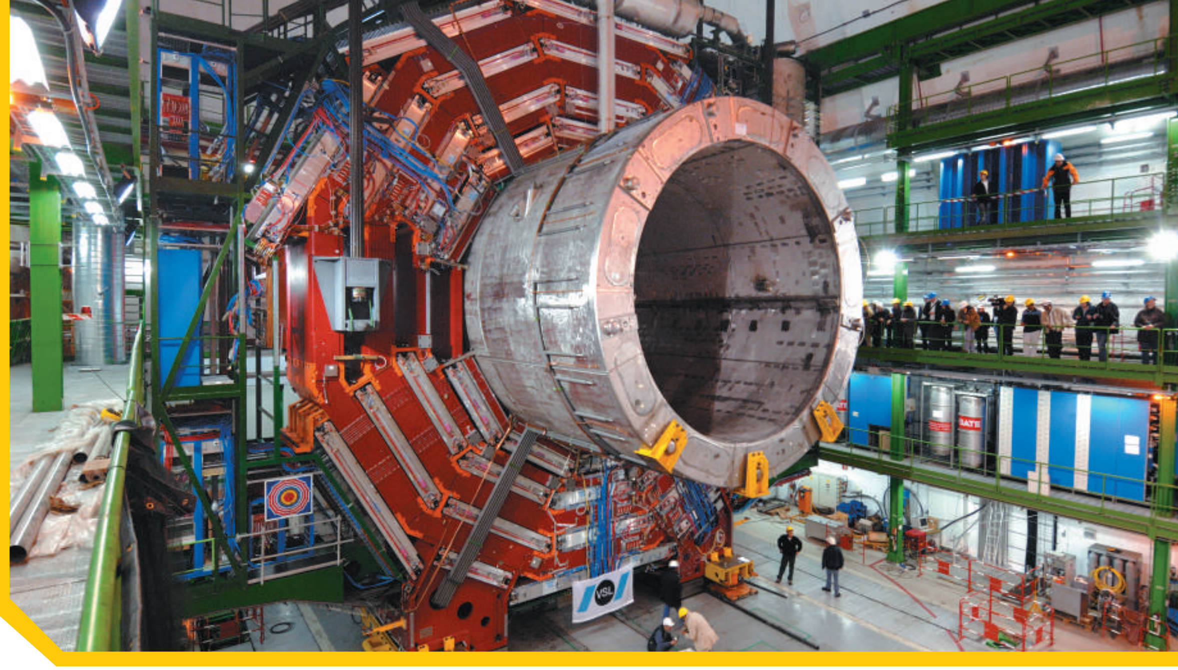
(ii) A magnetized compass needle has a magnetic dipole moment along its length, and the earth's magnetic field (which points generally northward) exerts a torque that tends to align that dipole moment with the field. See Section 27.7 for details.

Key Example VARIATION Problems

VP27.1.1 (a) 4.18×10^{-14} N (b) $-z$-direction
VP27.1.2 2.91×10^{-15} C
VP27.1.3 (a) 0.0175 T (b) $+z$-direction
VP27.1.4 (a) 0.189 T (b) $-y$-direction
VP27.6.1 (a) 0.435 mm (b) 2.87×10^7 rad/s (c) 4.57×10^6 Hz
VP27.6.2 (a) 0.348 mm (b) 1.64 mm (c) 4.80×10^{-16} N
VP27.6.3 (a) 8.00×10^5 m/s (b) $-y$-direction (c) $-x$-direction
VP27.6.4 (a) 6.64 cm (b) 7.47 cm
VP27.7.1 (a) 2.44×10^{-3} N (b) $+z$-direction
VP27.7.2 (a) 1.60 A (b) upward
VP27.7.3 (a) -2.52×10^{-3} N (b) -4.20×10^{-3} N (c) 0 N (d) 4.90×10^{-3} N
VP27.7.4 41.0° and 139.0°

Bridging Problem

(a) $\tau_x = -1.54 \times 10^{-4}$ N·m,
$\tau_y = -2.05 \times 10^{-4}$ N·m,
$\tau_z = -6.14 \times 10^{-4}$ N·m
(b) -7.55×10^{-4} J
(c) 42.1 rad/s



? The immense cylinder in this photograph is a current-carrying coil, or solenoid, that generates a uniform magnetic field in its interior as part of an experiment at CERN, the European Organization for Nuclear Research. If two such solenoids were joined end to end, the magnetic field along their common axis would (i) become four times stronger; (ii) double in strength; (iii) become $\sqrt{2}$ times stronger; (iv) not change; (v) weaken.

28 Sources of Magnetic Field

LEARNING OUTCOMES

In this chapter, you'll learn...

28.1 The nature of the magnetic field produced by a single moving charged particle.
28.2 How to describe the magnetic field produced by an element of a current-carrying conductor.
28.3 How to calculate the magnetic field produced by a long, straight, current-carrying wire.
28.4 Why wires carrying current in the same direction attract, while wires carrying opposing currents repel.
28.5 How to calculate the magnetic field produced by a current-carrying wire bent into a circle.
28.6 What Ampere's law is, and what it tells us about magnetic fields.
28.7 How to use Ampere's law to calculate the magnetic field of symmetric current distributions.
28.8 How microscopic currents within materials give them their magnetic properties.

You'll need to review...

10.5 Angular momentum of a particle.
21.3–21.5 Coulomb's law and electric-field calculations.
22.4 Solving problems with Gauss's law.
27.2–27.9 Magnetic field and magnetic force.

In Chapter 27 we studied the forces exerted on moving charges and on current-carrying conductors in a magnetic field. We didn't worry about how the magnetic field got there; we simply took its existence as a given fact. But how are magnetic fields *created*? We know that both permanent magnets and electric currents in electromagnets create magnetic fields. In this chapter we'll study these sources of magnetic field in detail.

We've learned that a charge creates an electric field and that an electric field exerts a force on a charge. But a *magnetic* field exerts a force on only a *moving* charge. Similarly, we'll see that only *moving* charges *create* magnetic fields. We'll begin our analysis with the magnetic field created by a single moving point charge. We can use this analysis to determine the field created by a small segment of a current-carrying conductor. Once we can do that, we can in principle find the magnetic field produced by *any* shape of conductor.

Then we'll introduce Ampere's law, which plays a role in magnetism analogous to the role of Gauss's law in electrostatics. Ampere's law lets us exploit symmetry properties in relating magnetic fields to their sources.

Moving charged particles within atoms respond to magnetic fields and can also act as sources of magnetic field. We'll use these ideas to understand how certain magnetic materials can be used to intensify magnetic fields as well as why some materials such as iron act as permanent magnets.

28.1 MAGNETIC FIELD OF A MOVING CHARGE

Let's start with the basics, the magnetic field of a single point charge q moving with a constant velocity $\vec{v}$. In practical applications, such as the solenoid shown in the photo that opens this chapter, magnetic fields are produced by tremendous numbers of charged particles moving together in a current. But once we understand how to calculate the magnetic field due to a single point charge, it's a small leap to calculate the field due to a current-carrying wire or collection of wires.

As we did for electric fields, we call the location of the moving charge at a given instant the **source point** and the point P where we want to find the field the **field point.** In Section 21.4 we found that at a field point a distance r from a point charge q, the magnitude of the *electric* field $\vec{E}$ caused by the charge is proportional to the charge magnitude $|q|$ and to $1/r^2$, and the direction of $\vec{E}$ (for positive q) is along the line from source point to field point. The corresponding relationship for the *magnetic* field $\vec{B}$ of a point charge q moving with constant velocity has some similarities and some interesting differences.

Experiments show that the magnitude of $\vec{B}$ is also proportional to $|q|$ and to $1/r^2$. But the *direction* of $\vec{B}$ is *not* along the line from source point to field point. Instead, $\vec{B}$ is perpendicular to the plane containing this line and the particle's velocity vector $\vec{v}$, as shown in **Fig. 28.1**. Furthermore, the field *magnitude* B is also proportional to the particle's speed v and to the sine of the angle ϕ. Thus the magnetic-field magnitude at point P is

$$B = \frac{\mu_0}{4\pi}\frac{|q|v\sin\phi}{r^2} \qquad (28.1)$$

The quantity μ_0 (read as "mu-nought" or "mu-sub-zero") is called the **magnetic constant.** The reason for including the factor of 4π will emerge shortly. We did something similar with Coulomb's law in Section 21.3.

Moving Charge: Vector Magnetic Field

We can incorporate both the magnitude and direction of $\vec{B}$ into a single vector equation by using the vector product. To avoid having to say "the direction from the source q to the field point P" over and over, we introduce a *unit* vector $\hat{r}$ ("r-hat") that points from the source point to the field point. (We used $\hat{r}$ for the same purpose in Section 21.4.) This unit vector is equal to the vector $\vec{r}$ from the source to the field point divided by its magnitude: $\hat{r} = \vec{r}/r$. Then

Magnetic field due to a point charge with constant velocity

Magnetic constant; Charge; Velocity; Unit vector from point charge toward where field is measured

$$\vec{B} = \frac{\mu_0}{4\pi}\frac{q\vec{v}\times\hat{r}}{r^2} \qquad (28.2)$$

Distance from point charge to where field is measured

Figure 28.1 shows the relationship of $\hat{r}$ to P and shows the magnetic field $\vec{B}$ at several points in the vicinity of the charge. At all points along a line through the charge parallel to the velocity $\vec{v}$, the field is zero because $\sin\phi = 0$ at all such points. At any distance r from q, $\vec{B}$ has its greatest magnitude at points lying in the plane perpendicular to $\vec{v}$, because there $\phi = 90°$ and $\sin\phi = 1$. If q is negative, the directions of $\vec{B}$ are opposite to those shown in Fig. 28.1.

Moving Charge: Magnetic Field Lines

A point charge in motion also produces an *electric* field, with field lines that radiate outward from a positive charge. The *magnetic* field lines are completely different. For a point charge moving with velocity $\vec{v}$, the magnetic field lines are *circles* centered on the line of $\vec{v}$ and lying in planes perpendicular to this line. The field-line directions for a positive charge are given by the following *right-hand rule,* one of several that we'll encounter in this chapter: Grasp the velocity vector $\vec{v}$ with your right hand so that your right thumb points in the direction of $\vec{v}$; your fingers then curl around the line of $\vec{v}$ in the same sense as the magnetic field lines, assuming q is positive. Figure 28.1a shows parts of a few field lines; Fig. 28.1b shows some field lines in a plane through q, perpendicular to $\vec{v}$. If the point charge is negative, the directions of the field and field lines are the opposite of those shown in Fig. 28.1.

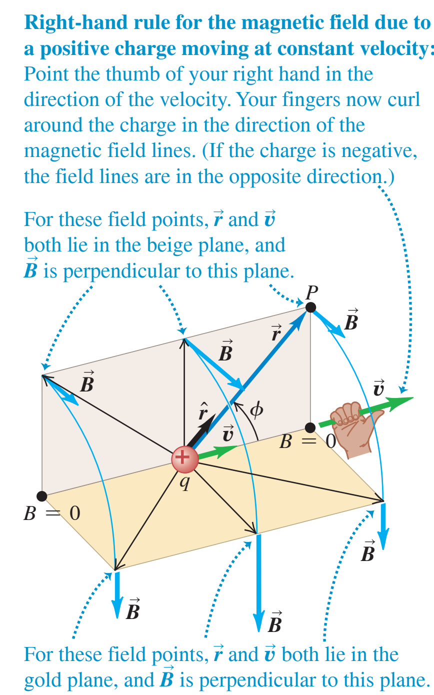
Equations (28.1) and (28.2) describe the $\vec{B}$ field of a point charge moving with *constant* velocity. If the charge *accelerates,* the field can be much more complicated. We won't need these more complicated results for our purposes. (The moving charged particles that make up a current in a wire accelerate at points where the wire bends and the direction of $\vec{v}$ changes. But because the magnitude v_d of the drift velocity in a conductor is typically very small, the centripetal acceleration v_d^2/r is so small that we can ignore its effects.)

As we discussed in Section 27.2, the unit of B is one tesla (1 T):

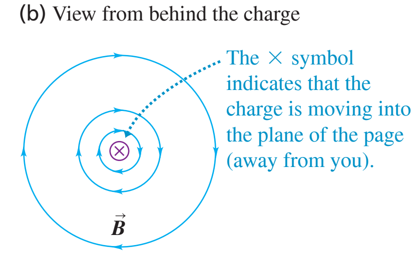
$$1\ \text{T} = 1\ \text{N}\cdot\text{s/C}\cdot\text{m} = 1\ \text{N/A}\cdot\text{m}$$

Figure **28.1** (a) Magnetic-field vectors due to a moving positive point charge q. At each point, $\vec{B}$ is perpendicular to the plane of $\vec{r}$ and $\vec{v}$, and its magnitude is proportional to the sine of the angle between them. (b) Magnetic field lines in a plane containing a moving positive charge.

(a) Perspective view



(b) View from behind the charge



Using this with Eq. (28.1) or (28.2), we find that the units of the constant μ_0 are

$$1\ \text{N}\cdot\text{s}^2/\text{C}^2 = 1\ \text{N/A}^2 = 1\ \text{Wb/A}\cdot\text{m} = 1\ \text{T}\cdot\text{m/A}$$

In SI units the numerical value of μ_0 is, to nine significant figures,

$$\begin{aligned}\mu_0 &= 1.25663706 \times 10^{-6}\ \text{N}\cdot\text{s}^2/\text{C}^2 = 1.25663706 \times 10^{-6}\ \text{Wb/A}\cdot\text{m}\\ &\cong 4\pi \times 10^{-7}\ \text{T}\cdot\text{m/A}\end{aligned} \tag{28.3}$$

It may seem incredible that μ_0 is equal to $4\pi \times 10^{-7}\ \text{T}\cdot\text{m/A}$ to nine significant figures! In fact this is a consequence of how the ampere and coulomb were previously defined. We'll explore this further in Section 28.4.

Here's another remarkable aspect of the value of the magnetic constant μ_0. Recall from Section 21.3 that the *electric* constant ϵ_0 that appears in Coulomb's law and Gauss's law has the value $\epsilon_0 = 8.854 \times 10^{-12}\ \text{C}^2/\text{N}\cdot\text{m}^2$. If we multiply ϵ_0 by $\mu_0 = 1.257 \times 10^{-6}\ \text{N}\cdot\text{s}^2/\text{C}^2$, take the square root of this product, then take the reciprocal of the result, we get

$$\frac{1}{\sqrt{\epsilon_0\mu_0}} = 2.998 \times 10^8\ \text{m/s}$$

But 2.998×10^8 m/s is just the value of the speed of light in vacuum! We conclude that

$$c = \frac{1}{\sqrt{\epsilon_0\mu_0}} \tag{28.4}$$

This suggests that electric and magnetic fields are intimately related to the nature of light. We'll explore this relationship when we study electromagnetic waves in Chapter 32.

EXAMPLE 28.1 Forces between two moving protons

WITH VARIATION PROBLEMS

Two protons move parallel to the x-axis in opposite directions (**Fig. 28.2**) at the same speed v (small compared to the speed of light c). At the instant shown, find the electric and magnetic forces on the upper proton and compare their magnitudes.

IDENTIFY and SET UP Coulomb's law [Eq. (21.2)] gives the electric force F_E on the upper proton. The magnetic force law [Eq. (27.2)] gives the magnetic force on the upper proton; to use it, we must first use Eq. (28.2) to find the magnetic field that the lower proton produces at the position of the upper proton. The unit vector from the lower proton (the source) to the position of the upper proton is $\hat{\boldsymbol{r}} = \hat{\boldsymbol{j}}$.

EXECUTE From Coulomb's law, the magnitude of the electric force on the upper proton is

$$F_E = \frac{1}{4\pi\epsilon_0}\frac{q^2}{r^2}$$

The forces are repulsive, and the force on the upper proton is vertically upward (in the $+y$-direction).

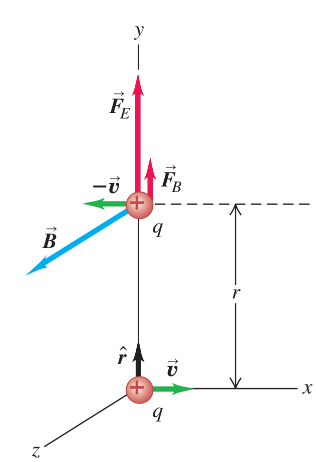
The velocity of the lower proton is $\vec{\boldsymbol{v}} = v\hat{\boldsymbol{\imath}}$. From the right-hand rule for the cross product $\vec{\boldsymbol{v}} \times \hat{\boldsymbol{r}}$ in Eq. (28.2), the $\vec{\boldsymbol{B}}$ field due to the lower proton at the position of the upper proton is in the $+z$-direction (see Fig. 28.2). From Eq. (28.2), the field is

$$\vec{\boldsymbol{B}} = \frac{\mu_0}{4\pi}\frac{q(v\hat{\boldsymbol{\imath}}) \times \hat{\boldsymbol{\jmath}}}{r^2} = \frac{\mu_0}{4\pi}\frac{qv}{r^2}\hat{\boldsymbol{k}}$$

The velocity of the upper proton is $-\vec{\boldsymbol{v}} = -v\hat{\boldsymbol{\imath}}$, so the magnetic force on it is

$$\vec{\boldsymbol{F}}_B = q(-\vec{\boldsymbol{v}}) \times \vec{\boldsymbol{B}} = q(-v\hat{\boldsymbol{\imath}}) \times \frac{\mu_0}{4\pi}\frac{qv}{r^2}\hat{\boldsymbol{k}} = \frac{\mu_0}{4\pi}\frac{q^2v^2}{r^2}\hat{\boldsymbol{\jmath}}$$

Figure **28.2** Electric and magnetic forces between two moving protons.



The magnetic interaction in this situation is also repulsive.

The ratio of the force magnitudes is

$$\frac{F_B}{F_E} = \frac{\mu_0 q^2v^2/4\pi r^2}{q^2/4\pi\epsilon_0 r^2} = \frac{\mu_0 v^2}{1/\epsilon_0} = \epsilon_0\mu_0 v^2$$

With the relationship $\epsilon_0\mu_0 = 1/c^2$, Eq. (28.4), this becomes

$$\frac{F_B}{F_E} = \frac{v^2}{c^2}$$

When v is small in comparison to the speed of light, the magnetic force is much smaller than the electric force.

EVALUATE We have described the velocities, fields, and forces as they are measured by an observer who is stationary in the coordinate system of Fig. 28.2. In a coordinate system that moves with one of the charges, one of the velocities would be zero, so there would be *no* magnetic force. The explanation of this apparent paradox provided one of the paths that led to the special theory of relativity.

KEYCONCEPT A moving charged particle produces a magnetic field $\vec{B}$. The magnitude of $\vec{B}$ at a given point is proportional to the charge magnitude and to the particle's speed, and inversely proportional to the square of the distance from the charge; it also depends on the direction from the particle to that point compared to the direction of the particle's velocity. The direction of $\vec{B}$ is given by a right-hand rule.

TEST YOUR UNDERSTANDING OF SECTION 28.1 (a) If two protons are traveling parallel to each other in the *same* direction and at the same speed, is the magnetic force between them (i) attractive or (ii) repulsive? (b) Is the net force between them (i) attractive, (ii) repulsive, or (iii) zero? (Assume that the protons' speed is much slower than the speed of light.)

ANSWER (a) **(i)**, (b) **(ii)** The situation is the same as shown in Fig. 28.2 except that the upper proton has velocity $\vec{v}$ rather than $-\vec{v}$. The magnetic field due to the lower proton is the same as shown in Fig. 28.2, but the direction of the magnetic force $\vec{F} = q\vec{v} \times \vec{B}$ on the upper proton is reversed. Hence the magnetic force is attractive. Since the speed v is small compared to c, the magnetic force is much smaller in magnitude than the repulsive electric force and the net force is still repulsive.

28.2 MAGNETIC FIELD OF A CURRENT ELEMENT

As for electric fields, there is a *principle of superposition of magnetic fields:*

PRINCIPLE OF SUPERPOSITION OF MAGNETIC FIELDS The total magnetic field caused by several moving charges is the vector sum of the fields caused by the individual charges.

We can use this principle with the results of Section 28.1 to find the magnetic field produced by a current in a conductor.

We begin by calculating the magnetic field caused by a short segment $d\vec{l}$ of a current-carrying conductor, as shown in **Fig. 28.3a**. The volume of the segment is $A\,dl$, where A is the cross-sectional area of the conductor. If there are n moving charged particles per unit volume, each of charge q, the total moving charge dQ in the segment is

$$dQ = nqA\,dl$$

The moving charges in this segment are equivalent to a single charge dQ, traveling with a velocity equal to the *drift* velocity $\vec{v}_d$. (Magnetic fields due to the *random* motions of the charges will, on average, cancel out at every point.) From Eq. (28.1) the magnitude of the resulting field $d\vec{B}$ at any field point P is

$$dB = \frac{\mu_0}{4\pi}\frac{|dQ|v_d \sin\phi}{r^2} = \frac{\mu_0}{4\pi}\frac{n|q|v_d A\,dl \sin\phi}{r^2}$$

But from Eq. (25.2), $n|q|v_d A$ equals the current I in the element. So

$$dB = \frac{\mu_0}{4\pi}\frac{I\,dl \sin\phi}{r^2} \tag{28.5}$$

Current Element: Vector Magnetic Field

In vector form, using the unit vector $\hat{r}$ as in Section 28.1, we have

Magnetic field due to an infinitesimal current element

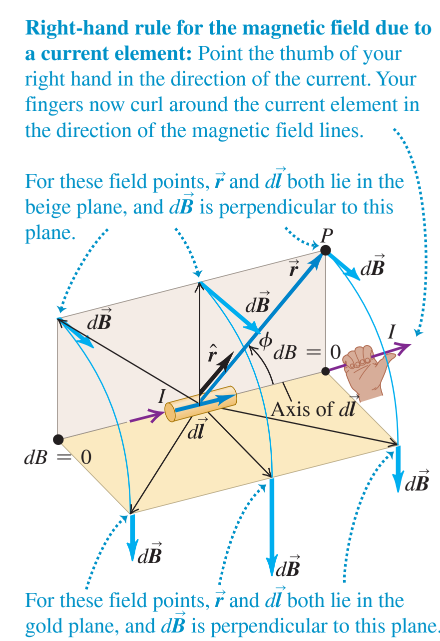
Magnetic constant; Current; Vector length of element (points in current direction); Unit vector from element toward where field is measured; Distance from element to where field is measured

$$d\vec{B} = \frac{\mu_0}{4\pi}\frac{I\,d\vec{l} \times \hat{r}}{r^2} \tag{28.6}$$

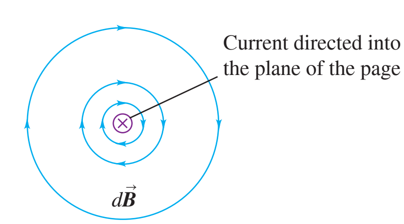
where $d\vec{l}$ is a vector with length dl, in the same direction as the current.

Figure **28.3** (a) Magnetic-field vectors due to a current element $d\vec{l}$. (b) Magnetic field lines in a plane containing the current element $d\vec{l}$. Compare this figure to Fig. 28.1 for the field of a moving point charge.

(a) Perspective view



(b) View along the axis of the current element



APPLICATION Currents and Planetary Magnetism The earth's magnetic field is caused by currents circulating within its molten, conducting interior. These currents are stirred by our planet's relatively rapid spin (one rotation per 24 hours). The planet Jupiter's internal currents are much stronger: Jupiter is much larger than the earth, has an interior that is mostly liquid hydrogen (which is a very good conductor under high pressures), and spins very rapidly (one rotation per 10 hours). Hence at the same distance from the center of each planet, Jupiter's magnetic field is about 2×10^4 times greater than the earth's field.

Equations (28.5) and (28.6) are called the **law of Biot and Savart** (pronounced "Bee-oh" and "Suh-var"). We can use this law to find the total magnetic field $\vec{B}$ at any point in space due to the current in a complete circuit. To do this, we integrate Eq. (28.6) over all segments $d\vec{l}$ that carry current; symbolically,

$$\vec{B} = \frac{\mu_0}{4\pi} \int \frac{I\, d\vec{l} \times \hat{r}}{r^2} \tag{28.7}$$

In the following sections we'll carry out this vector integration for several examples.

Current Element: Magnetic Field Lines

As Fig. 28.3 shows, the field vectors $d\vec{B}$ and the magnetic field lines of a current element are exactly like those set up by a positive charge dQ moving in the direction of the drift velocity $\vec{v}_d$. The field lines are circles in planes perpendicular to $d\vec{l}$ and centered on the line of $d\vec{l}$. Their directions are given by the same right-hand rule that we introduced for point charges in Section 28.1.

We can't verify Eq. (28.5) or (28.6) directly because we can never experiment with an isolated segment of a current-carrying circuit. What we measure experimentally is the *total* $\vec{B}$ for a complete circuit. But we can still verify these equations indirectly by calculating $\vec{B}$ for various current configurations with Eq. (28.7) and comparing the results with experimental measurements.

If matter is present in the space around a current-carrying conductor, the field at a field point P in its vicinity will have an additional contribution resulting from the *magnetization* of the material. We'll return to this point in Section 28.8. However, unless the material is iron or some other ferromagnetic material, the additional field is small and is usually negligible. Additional complications arise if time-varying electric or magnetic fields are present or if the material is a superconductor; we'll return to these topics later.

PROBLEM-SOLVING STRATEGY 28.1 Magnetic-Field Calculations

IDENTIFY *the relevant concepts:* The law of Biot and Savart [Eqs. (28.5) and (28.6)] allows you to calculate the magnetic field at a field point P due to a current-carrying wire of any shape. The idea is to calculate the field element $d\vec{B}$ at P due to a representative current element in the wire and integrate all such field elements to find the field $\vec{B}$ at P.

SET UP *the problem* using the following steps:

1. Make a diagram showing a representative current element and the field point P.
2. Draw the current element $d\vec{l}$, being careful that it points in the direction of the current.
3. Draw the unit vector $\hat{r}$ directed *from* the current element (the source point) to P.
4. Identify the target variable (usually $\vec{B}$).

EXECUTE *the solution* as follows:

1. Use Eq. (28.5) or (28.6) to express the magnetic field $d\vec{B}$ at P from the representative current element.
2. Add up all the $d\vec{B}$'s using the rules of *vector* addition to find the total field at point P. In some situations the $d\vec{B}$'s at point P have the same direction for all the current elements; then the magnitude of the total $\vec{B}$ field is the sum of the magnitudes of the $d\vec{B}$'s. But often the $d\vec{B}$'s have different directions for different current elements. Then you have to set up a coordinate system and represent each $d\vec{B}$ in terms of its components. The integral for the total $\vec{B}$ is then expressed in terms of an integral for each component.
3. Sometimes you can use the symmetry of the situation to prove that one component of $\vec{B}$ must vanish. Always be alert for ways to use symmetry to simplify the problem.
4. Look for ways to use the principle of superposition of magnetic fields. Later in this chapter we'll determine the fields produced by certain simple conductor shapes; if you encounter a conductor of a complex shape that can be represented as a combination of these simple shapes, you can use superposition to find the field of the complex shape. Examples include a rectangular loop and a semicircle with straight line segments on both sides.

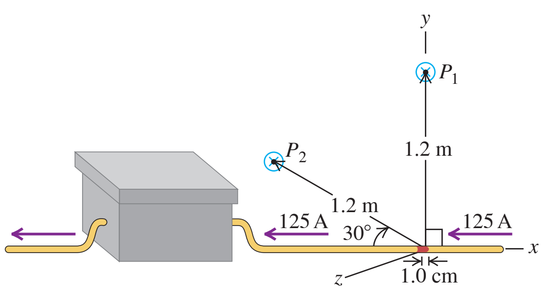
EVALUATE *your answer:* Often your answer will be a mathematical expression for $\vec{B}$ as a function of the position of the field point. Check the answer by examining its behavior in as many limits as you can.

EXAMPLE 28.2 Magnetic field of a current segment

WITH VARIATION PROBLEMS

A copper wire carries a steady 125 A current to an electroplating tank (**Fig. 28.4**). Find the magnetic field due to a 1.0 cm segment of this wire at a point 1.2 m away from it, if the point is (a) point P_1, straight out to the side of the segment, and (b) point P_2, in the xy-plane and on a line at 30° to the segment.

Figure **28.4** Finding the magnetic field at two points due to a 1.0 cm segment of current-carrying wire (not shown to scale).



IDENTIFY and SET UP Although Eqs. (28.5) and (28.6) apply only to infinitesimal current elements, we may use either of them here because the segment length is much less than the distance to the field point. The current element is shown in red in Fig. 28.4 and points in the $-x$-direction (the direction of the current), so $d\vec{l} = dl(-\hat{\imath})$. The unit vector $\hat{r}$ for each field point is directed from the current element toward that point: $\hat{r}$ is in the $+y$-direction for point P_1 and at an angle of 30° above the $-x$-direction for point P_2.

EXECUTE (a) At point P_1, $\hat{r} = \hat{\jmath}$, so

$$\vec{B} = \frac{\mu_0}{4\pi}\frac{I\,d\vec{l}\times\hat{r}}{r^2} = \frac{\mu_0}{4\pi}\frac{I\,dl(-\hat{\imath})\times\hat{\jmath}}{r^2} = -\frac{\mu_0}{4\pi}\frac{I\,dl}{r^2}\hat{k}$$

$$= -(10^{-7}\ \mathrm{T\cdot m/A})\frac{(125\ \mathrm{A})(1.0\times10^{-2}\ \mathrm{m})}{(1.2\ \mathrm{m})^2}\hat{k}$$

$$= -(8.7\times10^{-8}\ \mathrm{T})\hat{k}$$

The direction of $\vec{B}$ at P_1 is into the xy-plane of Fig. 28.4.

(b) At P_2, the unit vector is $\hat{r} = (-\cos 30°)\hat{\imath} + (\sin 30°)\hat{\jmath}$. From Eq. (28.6),

$$\vec{B} = \frac{\mu_0}{4\pi}\frac{I\,d\vec{l}\times\hat{r}}{r^2} = \frac{\mu_0}{4\pi}\frac{I\,dl(-\hat{\imath})\times(-\cos 30°\hat{\imath} + \sin 30°\hat{\jmath})}{r^2}$$

$$= -\frac{\mu_0 I}{4\pi}\frac{dl\sin 30°}{r^2}\hat{k}$$

$$= -(10^{-7}\ \mathrm{T\cdot m/A})\frac{(125\ \mathrm{A})(1.0\times10^{-2}\ \mathrm{m})(\sin 30°)}{(1.2\ \mathrm{m})^2}\hat{k}$$

$$= -(4.3\times10^{-8}\ \mathrm{T})\hat{k}$$

The direction of $\vec{B}$ at P_2 is also into the xy-plane of Fig. 28.4.

EVALUATE We can check our results for the direction of $\vec{B}$ by comparing them with Fig. 28.3. The xy-plane in Fig. 28.4 corresponds to the beige plane in Fig. 28.3, but here the direction of the current and hence of $d\vec{l}$ is the reverse of that shown in Fig. 28.3. Consequently the direction of the magnetic field is reversed as well. It follows that the field at points in the xy-plane in Fig. 28.4 must point *into,* not out of, that plane, just as we concluded above.

KEYCONCEPT The magnetic field $\vec{B}$ due to a short current segment of length dl is like that of a moving charged particle. The difference is that the product of the moving particle's charge and velocity is replaced by the product of the current and the vector $d\vec{l}$ in the direction of the current.

TEST YOUR UNDERSTANDING OF SECTION 28.2 An infinitesimal current element located at the origin $(x = y = z = 0)$ carries current I in the positive y-direction. Rank the following locations in order of the strength of the magnetic field that the current element produces at that location, from largest to smallest value. (i) $x = L, y = 0, z = 0$; (ii) $x = 0, y = L, z = 0$; (iii) $x = 0, y = 0, z = L$; (iv) $x = L/\sqrt{2}, y = L/\sqrt{2}, z = 0$.

ANSWER

(i) and (iii) (tie), (iv), (ii) From Eq. (28.5), the magnitude of the field dB due to a current element of length dl carrying current I is $dB = (\mu_0/4\pi)(I\,dl\sin\phi/r^2)$. In this expression r is the distance from the element to the field point, and ϕ is the angle between the direction of the current and a vector from the current element to the field point. All four points are the same distance $r = L$ from the current element, so the value of dB is proportional to the value of $\sin\phi$. For the four points the angle is (i) $\phi = 90°$, (ii) $\phi = 0$, (iii) $\phi = 90°$, and (iv) $\phi = 45°$, so the values of $\sin\phi$ are (i) 1, (ii) 0, (iii) 1, and (iv) $1/\sqrt{2}$.

28.3 MAGNETIC FIELD OF A STRAIGHT CURRENT-CARRYING CONDUCTOR

Let's use the law of Biot and Savart to find the magnetic field produced by a straight current-carrying conductor. This result is useful because straight conducting wires are found in essentially all electric and electronic devices. **Figure 28.5** (next page) shows such a conductor with length $2a$ carrying a current I. We'll find $\vec{B}$ at a point a distance x from the conductor on its perpendicular bisector.

We first use the law of Biot and Savart, Eq. (28.5), to find the field $d\vec{B}$ caused by the element of conductor of length $dl = dy$ shown in Fig. 28.5. From the figure, $r = \sqrt{x^2 + y^2}$ and $\sin\phi = \sin(\pi - \phi) = x/\sqrt{x^2 + y^2}$. The right-hand rule for the vector product $d\vec{l}\times\hat{r}$

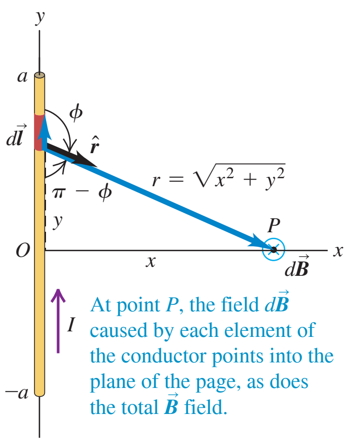


Figure 28.5 Magnetic field produced by a straight current-carrying conductor of length $2a$.

shows that the *direction* of $d\vec{B}$ is into the plane of the figure, perpendicular to the plane; furthermore, the directions of the $d\vec{B}$'s from *all* elements of the conductor are the same. Thus in integrating Eq. (28.7), we can just add the *magnitudes* of the $d\vec{B}$'s, a significant simplification.

Putting the pieces together, we find that the magnitude of the total $\vec{B}$ field is

$$B = \frac{\mu_0 I}{4\pi}\int_{-a}^{a}\frac{x\,dy}{(x^2+y^2)^{3/2}}$$

We can integrate this by trigonometric substitution or by using an integral table:

$$B = \frac{\mu_0 I}{4\pi}\frac{2a}{x\sqrt{x^2+a^2}} \tag{28.8}$$

When the length $2a$ of the conductor is much greater than its distance x from point P, we can consider it to be infinitely long. When a is much larger than x, $\sqrt{x^2 + a^2}$ is approximately equal to a; hence in the limit $a \to \infty$, Eq. (28.8) becomes

$$B = \frac{\mu_0 I}{2\pi x}$$

The physical situation has axial symmetry about the y-axis. Hence $\vec{B}$ must have the same *magnitude* at all points on a circle centered on the conductor and lying in a plane perpendicular to it, and the *direction* of $\vec{B}$ must be everywhere tangent to such a circle (**Fig. 28.6**). Thus, at all points on a circle of radius r around the conductor, the magnitude B is

Magnetic constant

Magnetic field near a long, straight, current-carrying conductor

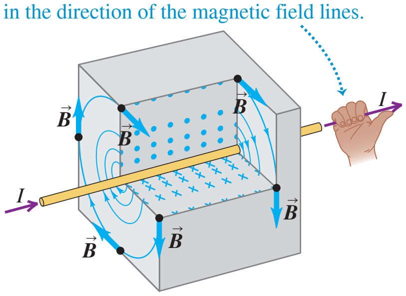
$$B = \frac{\mu_0 I}{2\pi r} \tag{28.9}$$

Current

Distance from conductor

Figure 28.6 Magnetic field around a long, straight, current-carrying conductor. The field lines are circles, with directions determined by the right-hand rule.

Right-hand rule for the magnetic field around a current-carrying wire: Point the thumb of your right hand in the direction of the current. Your fingers now curl around the wire in the direction of the magnetic field lines.



The geometry in this case is similar to that of Example 21.10 (Section 21.5), in which we solved the problem of the *electric* field caused by an infinite line of charge. The same integral appears in both problems, and the field magnitudes in both problems are proportional to $1/r$. But the lines of $\vec{B}$ in the magnetic problem have completely different shapes than the lines of $\vec{E}$ in the analogous electrical problem. Electric field lines radiate outward from a positive line charge distribution (inward for negative charges). By contrast, magnetic field lines *encircle* the current that acts as their source. Electric field lines due to charges begin and end at those charges, but magnetic field lines always form closed loops and *never* have endpoints, irrespective of the shape of the current-carrying conductor that sets up the field. As we discussed in Section 27.3, this is a consequence of Gauss's law for magnetism, which states that the total magnetic flux through *any* closed surface is always zero:

$$\oint \vec{B}\cdot d\vec{A} = 0 \quad \text{(magnetic flux through any closed surface)} \tag{28.10}$$

Any magnetic field line that enters a closed surface must emerge from that surface.

EXAMPLE 28.3 Magnetic field of a single wire

WITH VARIATION PROBLEMS

A long, straight conductor carries a 1.0 A current. At what distance from the axis of the conductor does the resulting magnetic field have magnitude $B = 0.5 \times 10^{-4}$ T (about that of the earth's magnetic field in Pittsburgh)?

IDENTIFY and SET UP The length of a "long" conductor is much greater than the distance from the conductor to the field point. Hence we can use the ideas of this section. The geometry is the same as that of Fig. 28.6, so we use Eq. (28.9). All of the quantities in this equation are known except the target variable, the distance r.

EXECUTE We solve Eq. (28.9) for r:

$$r = \frac{\mu_0 I}{2\pi B} = \frac{(4\pi \times 10^{-7}\ \text{T}\cdot\text{m/A})(1.0\ \text{A})}{(2\pi)(0.5 \times 10^{-4}\ \text{T})} = 4 \times 10^{-3}\ \text{m} = 4\ \text{mm}$$

EVALUATE As we saw in Example 26.14, currents of an ampere or more are typical of those found in the wiring of home appliances. This example shows that the magnetic fields produced by these appliances are very weak even very close to the wire; the fields are proportional to $1/r$, so they become even weaker at greater distances.

KEYCONCEPT The magnetic field $\vec{B}$ of a long, straight, current-carrying wire is proportional to the current and inversely proportional to the distance from the wire. The field lines are concentric circles around the wire, with the field direction given by a right-hand rule.

EXAMPLE 28.4 Magnetic field of two wires

WITH VARIATION PROBLEMS

Figure 28.7a is an end-on view of two long, straight, parallel wires perpendicular to the xy-plane, each carrying a current I but in opposite directions. (a) Find $\vec{B}$ at points P_1, P_2, and P_3. (b) Find an expression for $\vec{B}$ at any point on the x-axis to the right of wire 2.

IDENTIFY and SET UP We can find the magnetic fields $\vec{B}_1$ and $\vec{B}_2$ due to wires 1 and 2 at each point by using the ideas of this section. By the superposition principle, the magnetic field at each point is then $\vec{B} = \vec{B}_1 + \vec{B}_2$. We use Eq. (28.9) to find the magnitudes B_1 and B_2 of these fields and the right-hand rule to find the corresponding directions. Figure 28.7a shows $\vec{B}_1$, $\vec{B}_2$, and $\vec{B} = \vec{B}_{\text{total}}$ at each point; you should confirm that the directions and relative magnitudes shown are correct. Figure 28.7b shows some of the magnetic field lines due to this two-wire system.

EXECUTE (a) Since point P_1 is a distance $2d$ from wire 1 and a distance $4d$ from wire 2, $B_1 = \mu_0 I/2\pi(2d) = \mu_0 I/4\pi d$ and $B_2 = \mu_0 I/2\pi(4d) = \mu_0 I/8\pi d$. The right-hand rule shows that $\vec{B}_1$ is in the negative y-direction and $\vec{B}_2$ is in the positive y-direction, so

$$\vec{B}_{\text{total}} = \vec{B}_1 + \vec{B}_2 = -\frac{\mu_0 I}{4\pi d}\hat{\jmath} + \frac{\mu_0 I}{8\pi d}\hat{\jmath} = -\frac{\mu_0 I}{8\pi d}\hat{\jmath} \qquad \text{(point } P_1\text{)}$$

At point P_2, a distance d from both wires, $\vec{B}_1$ and $\vec{B}_2$ are both in the positive y-direction, and both have the same magnitude $B_1 = B_2 = \mu_0 I/2\pi d$. Hence

$$\vec{B}_{\text{total}} = \vec{B}_1 + \vec{B}_2 = \frac{\mu_0 I}{2\pi d}\hat{\jmath} + \frac{\mu_0 I}{2\pi d}\hat{\jmath} = \frac{\mu_0 I}{\pi d}\hat{\jmath} \qquad \text{(point } P_2\text{)}$$

Finally, at point P_3 the right-hand rule shows that $\vec{B}_1$ is in the positive y-direction and $\vec{B}_2$ is in the negative y-direction. This point is a distance $3d$ from wire 1 and a distance d from wire 2, so $B_1 = \mu_0 I/2\pi(3d) = \mu_0 I/6\pi d$ and $B_2 = \mu_0 I/2\pi d$. The total field at P_3 is

$$\vec{B}_{\text{total}} = \vec{B}_1 + \vec{B}_2 = \frac{\mu_0 I}{6\pi d}\hat{\jmath} - \frac{\mu_0 I}{2\pi d}\hat{\jmath} = -\frac{\mu_0 I}{3\pi d}\hat{\jmath} \qquad \text{(point } P_3\text{)}$$

The same technique can be used to find $\vec{B}_{\text{total}}$ at any point; for points off the x-axis, caution must be taken in vector addition, since $\vec{B}_1$ and $\vec{B}_2$ need no longer be simply parallel or antiparallel.

(b) At any point on the x-axis to the right of wire 2 (that is, for $x > d$), $\vec{B}_1$ and $\vec{B}_2$ are in the same directions as at P_3. Such a point is a distance $x + d$ from wire 1 and a distance $x - d$ from wire 2, so the total field is

$$\vec{B}_{\text{total}} = \vec{B}_1 + \vec{B}_2 = \frac{\mu_0 I}{2\pi(x+d)}\hat{\jmath} - \frac{\mu_0 I}{2\pi(x-d)}\hat{\jmath}$$

$$= -\frac{\mu_0 I d}{\pi(x^2 - d^2)}\hat{\jmath}$$

where we used a common denominator to combine the two terms.

EVALUATE Consider our result from part (b) at a point very far from the wires, so that x is much larger than d. Then the d^2 term in the denominator can be ignored, and the magnitude of the total field is approximately $B_{\text{total}} = \mu_0 I d/\pi x^2$. For one wire, Eq. (28.9) shows that the magnetic field decreases with distance in proportion to $1/x$; for two wires carrying opposite currents, $\vec{B}_1$ and $\vec{B}_2$ partially cancel each other, and so B_{total} decreases more rapidly, in proportion to $1/x^2$. This effect is used in communication systems such as telephone or computer networks. The wiring is arranged so that a conductor carrying a signal in one direction and the conductor carrying the return signal are side by side, as in Fig. 28.7a, or twisted around each other (**Fig. 28.8**). As a result, the magnetic field due to these signals *outside* the conductors is very small, making it less likely to exert unwanted forces on other information-carrying currents.

Figure **28.7** (a) Two long, straight conductors carrying equal currents in opposite directions. The conductors are seen end-on. (b) Map of the magnetic field produced by the two conductors. The field lines are closest together between the conductors, where the field is strongest.

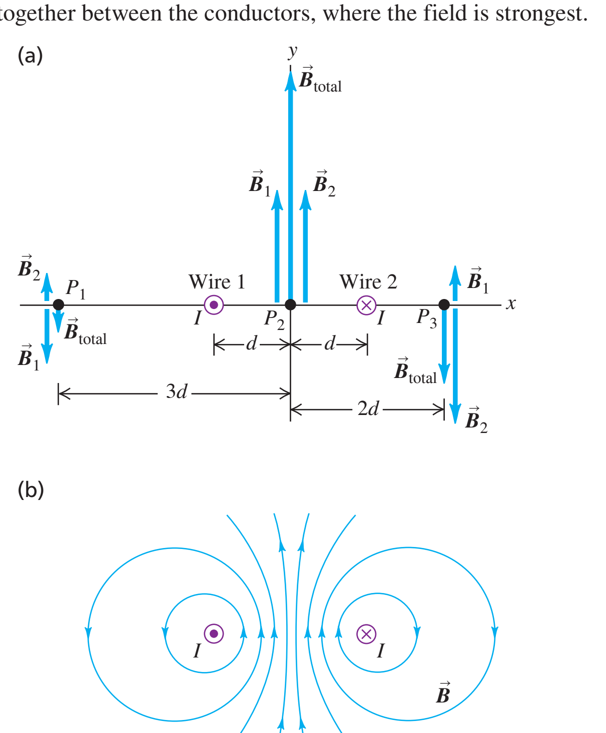
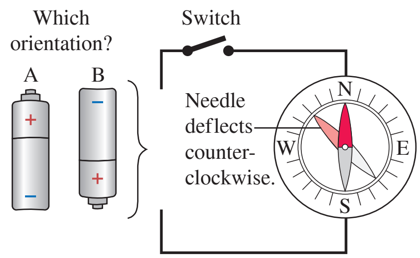


Figure **28.8** Computer cables, or cables for audio-video equipment, create little or no magnetic field. This is because within each cable, closely spaced wires carry current in both directions along the length of the cable. The magnetic fields from these opposing currents cancel each other.

KEYCONCEPT Magnetic fields obey the superposition principle: The net magnetic field due to two or more sources is the vector sum of the fields due to the individual sources.



TEST YOUR UNDERSTANDING OF SECTION 28.3 The accompanying figure shows a circuit that lies on a horizontal table. A compass is placed on top of the circuit as shown. A battery is to be connected to the circuit so that when the switch is closed, the compass needle deflects counterclockwise. In which orientation, A or B, should the battery be placed in the circuit?

ANSWER

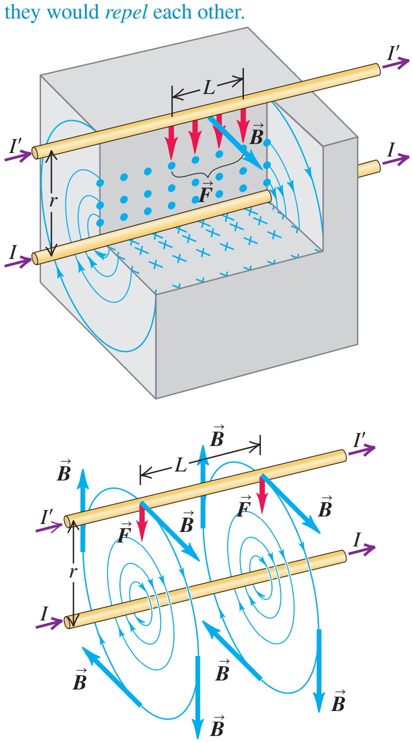
❙ A This orientation will cause current to flow clockwise around the circuit. Hence current will flow south through the wire that lies under the compass. From the right-hand rule for the magnetic field produced by a long, straight, current-carrying conductor, this will produce a magnetic field that points to the left at the position of the compass (which lies atop the wire). The combination of the northward magnetic field of the earth and the westward field produced by the current gives a net magnetic field to the northwest, so the compass needle will swing counterclockwise to align with this field.

28.4 FORCE BETWEEN PARALLEL CONDUCTORS

Figure **28.9** Parallel conductors carrying currents in the same direction attract each other. The diagrams show how the magnetic field $\vec{B}$ caused by the current in the lower conductor exerts a force $\vec{F}$ on the upper conductor.

The magnetic field of the lower wire exerts an attractive force on the upper wire. By the same token, the upper wire attracts the lower one.

If the wires had currents in *opposite* directions, they would *repel* each other.



Now that we know how to calculate the magnetic field produced by a long, current-carrying conductor, we can find the *magnetic force* that one such conductor exerts on another. This force plays a role in many practical situations in which current-carrying wires are close to each other. **Figure 28.9** shows segments of two long, straight, parallel conductors separated by a distance r and carrying currents I and I' in the same direction. Each conductor lies in the magnetic field set up by the other, so each experiences a force. The figure shows some of the field lines set up by the current in the lower conductor.

From Eq. (28.9) the lower conductor produces a $\vec{B}$ field that, at the position of the upper conductor, has magnitude

$$B = \frac{\mu_0 I}{2\pi r}$$

From Eq. (27.19) the force that this field exerts on a length L of the upper conductor is $\vec{F} = I'\vec{L} \times \vec{B}$, where the vector $\vec{L}$ is in the direction of the current I' and has magnitude L. Since $\vec{B}$ is perpendicular to the length of the conductor and hence to $\vec{L}$, the magnitude of this force is

$$F = I'LB = \frac{\mu_0 I I' L}{2\pi r}$$

and the force *per unit length* F/L is

Magnetic force per unit length between two long, straight, parallel, current-carrying conductors

$$\frac{F}{L} = \frac{\mu_0 I I'}{2\pi r} \tag{28.11}$$

Magnetic constant; Current in first conductor; Current in second conductor; Distance between conductors

Applying the right-hand rule to $\vec{F} = I'\vec{L} \times \vec{B}$ shows that the force on the upper conductor is directed *downward.*

The current in the *upper* conductor also sets up a $\vec{B}$ field at the position of the *lower* conductor. Two successive applications of the right-hand rule for vector products (one to find the direction of the $\vec{B}$ field due to the upper conductor, as in Section 28.2, and one to find the direction of the force that this field exerts on the lower conductor, as in Section 27.6) show that the force on the lower conductor is *upward.* Thus *two parallel conductors carrying current in the same direction attract each other.* If the direction of either current is reversed, the forces also reverse. *Parallel conductors carrying currents in opposite directions repel each other.*

Magnetic Forces and the Value of μ_0

We saw in Section 28.1 that the magnetic constant μ_0 is very nearly equal to $4\pi \times 10^{-7}\ \mathrm{T \cdot m/A}$. If μ_0 had *exactly* this value, Eq. (28.11) shows that the force per unit length on each of two infinitely long parallel conductors one meter apart, each carrying a current of one ampere, would be

$$\frac{F}{L} = \frac{\mu_0 I I'}{2\pi r} = \frac{(4\pi \times 10^{-7}\ \mathrm{T \cdot m/A})(1\ \mathrm{A})(1\ \mathrm{A})}{2\pi(1\ \mathrm{m})} = 2 \times 10^{-7}\ \mathrm{T \cdot A} = 2 \times 10^{-7}\ \mathrm{N/m}$$

Prior to 2018 the ampere was defined so that F/L would have exactly this value in this situation, and μ_0 was indeed exactly equal to $4\pi \times 10^{-7}\ \mathrm{T \cdot m/A}$. The coulomb was defined to be the amount of charge transferred in one second by a current of one ampere.

As we described in Section 21.3, the coulomb is now defined in terms of e, the magnitude of charge on an electron or proton. Hence the value of μ_0 has changed very slightly from $4\pi \times 10^{-7}\ \mathrm{T \cdot m/A}$ and has a fractional uncertainty of about 2×10^{-10}. The difference between the old and new values of μ_0 is so miniscule, however, that we can safely use $4\pi \times 10^{-7}\ \mathrm{T \cdot m/A}$ in calculations.

Mutual forces of attraction exist not only between *wires* carrying currents in the same direction, but also between the elements of a single current-carrying conductor. If the conductor is a liquid or an ionized gas (a plasma), these forces result in a constriction of the conductor called the *pinch effect*. The pinch effect in a plasma has been used in one technique to bring about nuclear fusion.

EXAMPLE 28.5 Forces between parallel wires

WITH VARIATION PROBLEMS

Two straight, parallel, superconducting wires 4.5 mm apart carry equal currents of 15,000 A in opposite directions. What force, per unit length, does each wire exert on the other?

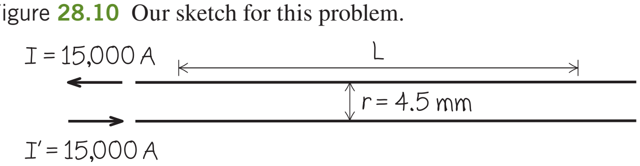
IDENTIFY and SET UP **Figure 28.10** shows the situation. We find F/L, the magnetic force per unit length of wire, from Eq. (28.11).

EXECUTE The conductors *repel* each other because the currents are in opposite directions. From Eq. (28.11) the force per unit length is

$$\frac{F}{L} = \frac{\mu_0 I I'}{2\pi r} = \frac{(4\pi \times 10^{-7}\ \mathrm{T \cdot m/A})(15{,}000\ \mathrm{A})^2}{(2\pi)(4.5 \times 10^{-3}\ \mathrm{m})}$$

$$= 1.0 \times 10^4\ \mathrm{N/m}$$

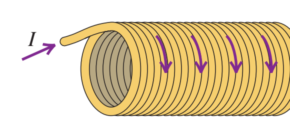
Figure 28.10 Our sketch for this problem.



EVALUATE This is a large force, more than one ton per meter. Currents and separations of this magnitude are used in superconducting electromagnets in particle accelerators, and mechanical stress analysis is a crucial part of the design process.

KEYCONCEPT Two long, straight, parallel, current-carrying conductors attract each other if they carry current in the same direction. They repel each other if they carry current in opposite directions.

TEST YOUR UNDERSTANDING OF SECTION 28.4 A solenoid is a wire wound into a helical coil. The accompanying figure shows a solenoid that carries a current I. (a) Is the *magnetic* force that one turn of the coil exerts on an adjacent turn (i) attractive, (ii) repulsive, or (iii) zero? (b) Is the *electric* force that one turn of the coil exerts on an adjacent turn (i) attractive, (ii) repulsive, or (iii) zero? (c) Is the *magnetic* force between opposite sides of the same turn of the coil (i) attractive, (ii) repulsive, or (iii) zero? (d) Is the *electric* force between opposite sides of the same turn of the coil (i) attractive, (ii) repulsive, or (iii) zero?



ANSWER

(a) **(i)**, (b) **(iii)**, (c) **(ii)**, (d) **(iii)** Current flows in the same direction in adjacent turns of the coil, so the magnetic forces between these turns are attractive. Current flows in opposite directions on opposite sides of the same turn, so the magnetic forces between these sides are repulsive. Thus the magnetic forces on the solenoid turns squeeze them together in the direction along its axis but push them apart radially. The *electric* forces are zero because the wire is electrically neutral, with as much positive charge as there is negative charge.

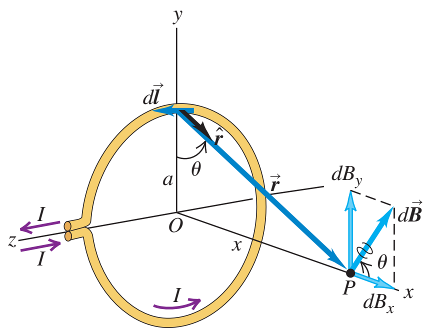
28.5 MAGNETIC FIELD OF A CIRCULAR CURRENT LOOP

If you look inside a doorbell, a transformer, an electric motor, or an electromagnet (**Fig. 28.11**), you'll find coils of wire with a large number of turns, spaced so closely that each turn is very nearly a planar circular loop. A current in such a coil is used to establish a magnetic field. In Section 27.7 we considered the force and torque on such a current loop placed in an external magnetic field produced by other currents; we are now about to find the magnetic field produced by such a loop or by a collection of closely spaced loops forming a coil.

Figure 28.12 (next page) shows a circular conductor with radius a. A current I is led into and out of the loop through two long, straight wires side by side; the currents in these straight wires are in opposite directions, and their magnetic fields very nearly cancel each other (see Example 28.4 in Section 28.3).

Figure 28.11 This electromagnet contains a current-carrying coil with numerous turns of wire. The resulting magnetic field can pick up large quantities of steel bars and other iron-bearing items.

Figure 28.12 Magnetic field on the axis of a circular loop. The current in the segment $d\vec{l}$ causes the field $d\vec{B}$, which lies in the xy-plane. The currents in other $d\vec{l}$'s cause $d\vec{B}$'s with different components perpendicular to the x-axis; these components add to zero. The x-components of the $d\vec{B}$'s combine to give the total $\vec{B}$ field at point P.



We can use the law of Biot and Savart, Eq. (28.5) or (28.6), to find the magnetic field at a point P on the axis of the loop, at a distance x from the center. As the figure shows, $d\vec{l}$ and $\hat{r}$ are perpendicular, and the direction of the field $d\vec{B}$ caused by this particular element $d\vec{l}$ lies in the xy-plane. Since $r^2 = x^2 + a^2$, the magnitude dB of the field due to element $d\vec{l}$ is

$$dB = \frac{\mu_0 I}{4\pi} \frac{dl}{(x^2 + a^2)} \tag{28.12}$$

The components of the vector $d\vec{B}$ are

$$dB_x = dB\cos\theta = \frac{\mu_0 I}{4\pi} \frac{dl}{(x^2 + a^2)} \frac{a}{(x^2 + a^2)^{1/2}} \tag{28.13}$$

$$dB_y = dB\sin\theta = \frac{\mu_0 I}{4\pi} \frac{dl}{(x^2 + a^2)} \frac{x}{(x^2 + a^2)^{1/2}} \tag{28.14}$$

The *total* field $\vec{B}$ at P has only an x-component (it is perpendicular to the plane of the loop). Here's why: For every element $d\vec{l}$ there is a corresponding element on the opposite side of the loop, with opposite direction. These two elements give equal contributions to the x-component of $d\vec{B}$, given by Eq. (28.13), but *opposite* components perpendicular to the x-axis. Thus all the perpendicular components cancel and only the x-components survive.

To obtain the x-component of the total field $\vec{B}$, we integrate Eq. (28.13), including all the $d\vec{l}$'s around the loop. Everything in this expression except dl is constant and can be taken outside the integral, and we have

$$B_x = \int \frac{\mu_0 I}{4\pi} \frac{a\,dl}{(x^2 + a^2)^{3/2}} = \frac{\mu_0 I a}{4\pi(x^2 + a^2)^{3/2}} \int dl$$

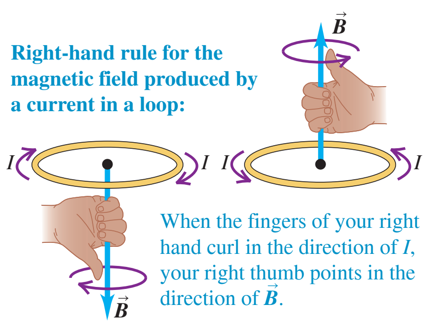
The integral of dl is just the circumference of the circle, $\int dl = 2\pi a$, and so

Magnetic field on axis of a circular current-carrying loop

Magnetic constant; Current; Radius of loop; Distance along axis from center of loop to field point

$$B_x = \frac{\mu_0 I a^2}{2(x^2 + a^2)^{3/2}} \tag{28.15}$$

Figure 28.13 The right-hand rule for the direction of the magnetic field produced on the axis of a current-carrying coil.



The *direction* of this magnetic field is given by a right-hand rule. If you curl the fingers of your right hand around the loop in the direction of the current, your right thumb points in the direction of the field (**Fig. 28.13**).

Magnetic Field on the Axis of a Coil

Now suppose that instead of the single loop in Fig. 28.12 we have a coil consisting of N loops, all with the same radius. The loops are closely spaced so that the plane of each loop is essentially the same distance x from the field point P. Then the total field is N times the field of a single loop:

$$B_x = \frac{\mu_0 N I a^2}{2(x^2 + a^2)^{3/2}} \quad \text{(on the axis of } N \text{ circular loops)} \tag{28.16}$$

The factor N in Eq. (28.16) is the reason coils of wire, not single loops, are used to produce strong magnetic fields; for a desired field strength, using a single loop might require a current I so great as to exceed the rating of the loop's wire.

Figure 28.14 Graph of the magnetic field along the axis of a circular coil with N turns. When x is much larger than a, the field magnitude decreases approximately as $1/x^3$.

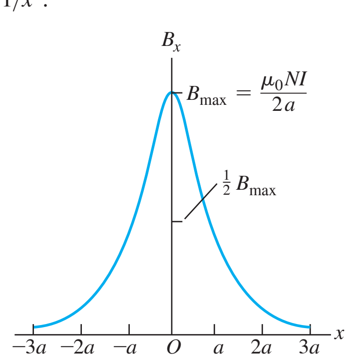


Figure 28.14 shows a graph of B_x as a function of x. The maximum value of the field is at $x = 0$, the center of the loop or coil:

Magnetic field at center of N circular current-carrying loops

Magnetic constant; Number of loops; Current; Radius of loop

$$B_x = \frac{\mu_0 N I}{2a} \tag{28.17}$$

In Section 27.7 we defined the *magnetic dipole moment* μ (or *magnetic moment*) of a current-carrying loop to be equal to IA, where A is the cross-sectional area of the loop. If there are N loops, the total magnetic moment is NIA. The circular loop in Fig. 28.12 has area $A = \pi a^2$, so the magnetic moment of a single loop is $\mu = I\pi a^2$; for N loops, $\mu = NI\pi a^2$. Substituting these results into Eqs. (28.15) and (28.16), we find

$$B_x = \frac{\mu_0 \mu}{2\pi(x^2 + a^2)^{3/2}} \quad \text{(on the axis of any number of circular loops)} \qquad (28.18)$$

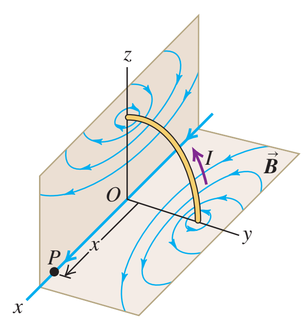
CAUTION Magnetic field of a coil Equations (28.15), (28.16), and (28.18) are valid only on the *axis* of a loop or coil. Don't attempt to apply these equations at other points!

We described a magnetic dipole in Section 27.7 in terms of its response to a magnetic field produced by currents outside the dipole. But a magnetic dipole is also a *source* of magnetic field; Eq. (28.18) describes the magnetic field *produced* by a magnetic dipole for points along the dipole axis. This field is directly proportional to the magnetic dipole moment μ. Note that the field at all points on the x-axis is in the same direction as the vector magnetic moment $\vec{\boldsymbol{\mu}}$.

Figure 28.15 shows some of the magnetic field lines surrounding a circular current loop (magnetic dipole) in planes through the axis. The directions of the field lines are given by the same right-hand rule as for a long, straight conductor. Grab the conductor with your right hand, with your thumb in the direction of the current; your fingers curl around in the same direction as the field lines. The field lines for the circular current loop are closed curves that encircle the conductor; they are *not* circles, however.

BIO APPLICATION Magnetic Fields for MRI Magnetic resonance imaging (see Section 27.7) requires a magnetic field of about 1.5 T. In a typical MRI scan, the patient lies inside a coil that produces the intense field. The currents required are very high, so the coils are bathed in liquid helium at a temperature of 4.2 K to keep them from overheating.

Figure **28.15** Magnetic field lines produced by the current in a circular loop. At points on the axis the $\vec{B}$ field has the same direction as the magnetic moment of the loop.



EXAMPLE 28.6 Magnetic field of a coil

A coil consisting of 100 circular loops with radius 0.60 m carries a 5.0 A current. (a) Find the magnetic field at a point along the axis of the coil, 0.80 m from the center. (b) Along the axis, at what distance from the center of the coil is the field magnitude $\frac{1}{8}$ as great as it is at the center?

IDENTIFY and SET UP This problem concerns the magnetic-field magnitude B along the axis of a current-carrying coil, so we can use Eq. (28.16). We are given $N = 100$, $I = 5.0$ A, and $a = 0.60$ m. In part (a) our target variable is B_x at a given value of x. In part (b) the target variable is the value of x at which the field has $\frac{1}{8}$ of the magnitude that it has at the origin.

EXECUTE (a) Using $x = 0.80$ m, from Eq. (28.16) we have

$$B_x = \frac{(4\pi \times 10^{-7}\ \mathrm{T \cdot m/A})(100)(5.0\ \mathrm{A})(0.60\ \mathrm{m})^2}{2[(0.80\ \mathrm{m})^2 + (0.60\ \mathrm{m})^2]^{3/2}}$$
$$= 1.1 \times 10^{-4}\ \mathrm{T}$$

(b) Considering Eq. (28.16), we seek a value of x such that

$$\frac{1}{(x^2 + a^2)^{3/2}} = \frac{1}{8}\frac{1}{(0^2 + a^2)^{3/2}}$$

To solve this for x, we take the reciprocal of the whole thing and then take the $\frac{2}{3}$ power of both sides; the result is

$$x = \pm\sqrt{3}a = \pm 1.04\ \mathrm{m}$$

EVALUATE We check our answer in part (a) by finding the coil's magnetic moment and substituting the result into Eq. (28.18):

$$\mu = NI\pi a^2 = (100)(5.0\ \mathrm{A})\pi(0.60\ \mathrm{m})^2 = 5.7 \times 10^2\ \mathrm{A \cdot m^2}$$

$$B_x = \frac{(4\pi \times 10^{-7}\ \mathrm{T \cdot m/A})(5.7 \times 10^2\ \mathrm{A \cdot m^2})}{2\pi[(0.80\ \mathrm{m})^2 + (0.60\ \mathrm{m})^2]^{3/2}} = 1.1 \times 10^{-4}\ \mathrm{T}$$

The magnetic moment μ is relatively large, yet it produces a rather small field, comparable to that of the earth. This illustrates how difficult it is to produce strong fields of 1 T or more.

KEYCONCEPT The magnetic field $\vec{B}$ on the axis of a current-carrying coil points along the axis, with a direction given by a right-hand rule. The magnitude of $\vec{B}$ is proportional to the magnetic dipole moment of the coil and is greatest at the center of the coil.

TEST YOUR UNDERSTANDING OF SECTION 28.5 Figure 28.12 shows the magnetic field $d\vec{B}$ produced at point P by a segment $d\vec{l}$ that lies on the positive y-axis (at the top of the loop). This field has components $dB_x > 0$, $dB_y > 0$, $dB_z = 0$. (a) What are the signs of the components of the field $d\vec{B}$ produced at P by a segment $d\vec{l}$ on the negative y-axis (at the bottom of the loop)? (i) $dB_x > 0$, $dB_y > 0$, $dB_z = 0$; (ii) $dB_x > 0$, $dB_y < 0$, $dB_z = 0$; (iii) $dB_x < 0$, $dB_y > 0$, $dB_z = 0$; (iv) $dB_x < 0$, $dB_y < 0$, $dB_z = 0$; (v) none of these. (b) What are the signs of the components of the field $d\vec{B}$ produced at P by a segment $d\vec{l}$ on the negative z-axis (at the right-hand side of the loop)? (i) $dB_x > 0$, $dB_y > 0$, $dB_z = 0$; (ii) $dB_x > 0$, $dB_y < 0$, $dB_z = 0$; (iii) $dB_x < 0$, $dB_y > 0$, $dB_z = 0$; (iv) $dB_x < 0$, $dB_y < 0$, $dB_z = 0$; (v) none of these.

ANSWER

(a) **(ii)**, (b) **(v)** The vector $d\vec{B}$ is in the direction of $d\vec{l} \times \vec{r}$. For a segment on the negative y-axis, $d\vec{l} = -\hat{k}\,dl$ points in the negative z-direction and $\vec{r} = x\hat{\imath} + a\hat{\jmath}$. Hence $d\vec{l} \times \vec{r} = (a\,dl)\hat{\imath} - (x\,dl)\hat{\jmath}$, which has a positive x-component, a negative y-component, and zero z-component. For a segment on the negative z-axis, $d\vec{l} = \hat{\jmath}\,dl$ points in the positive y-direction and $\vec{r} = x\hat{\imath} + a\hat{k}$. Hence $d\vec{l} \times \vec{r} = (a\,dl)\hat{\imath} - (x\,dl)\hat{k}$, which has a positive x-component, zero y-component, and a negative z-component.

28.6 AMPERE'S LAW

So far our calculations of the magnetic field due to a current have involved finding the infinitesimal field $d\vec{B}$ due to a current element and then summing all the $d\vec{B}$'s to find the total field. This approach is directly analogous to our *electric*-field calculations in Chapter 21.

For the electric-field problem we found that in situations with a highly symmetric charge distribution, it was often easier to use Gauss's law to find $\vec{E}$. There is likewise a law that allows us to more easily find the *magnetic* fields caused by highly symmetric *current* distributions. But the law that allows us to do this, called *Ampere's law,* is rather different in character from Gauss's law.

Gauss's law for electric fields (Chapter 22) involves the flux of $\vec{E}$ through a closed surface; it states that this flux is equal to the total charge enclosed within the surface, divided by the constant ϵ_0. Thus this law relates electric fields and charge distributions. By contrast, Gauss's law for *magnetic* fields, Eq. (28.10), is *not* a relationship between magnetic fields and current distributions; it states that the flux of $\vec{B}$ through *any* closed surface is always zero, whether or not there are currents within the surface. So Gauss's law for $\vec{B}$ can't be used to determine the magnetic field produced by a particular current distribution.

Ampere's law is formulated not in terms of magnetic flux, but rather in terms of the *line integral* of $\vec{B}$ around a closed path, denoted by

$$\oint \vec{B} \cdot d\vec{l}$$

We used line integrals to define work in Chapter 6 and to calculate electric potential in Chapter 23. To evaluate this integral, we divide the path into infinitesimal segments $d\vec{l}$, calculate the scalar product of $\vec{B} \cdot d\vec{l}$ for each segment, and sum these products. In general, $\vec{B}$ varies from point to point, and we must use the value of $\vec{B}$ at the location of each $d\vec{l}$. An alternative notation is $\oint B_{\parallel}\,dl$, where $B_{\parallel}$ is the component of $\vec{B}$ parallel to $d\vec{l}$ at each point. The circle on the integral sign indicates that this integral is always computed for a *closed* path, one whose beginning and end points are the same.

Ampere's Law for a Long, Straight Conductor

To introduce the basic idea of Ampere's law, let's consider again the magnetic field caused by a long, straight conductor carrying a current I. We found in Section 28.3 that the field at a distance r from the conductor has magnitude

$$B = \frac{\mu_0 I}{2\pi r}$$

The magnetic field lines are circles centered on the conductor. Let's take the line integral of $\vec{B}$ around a circle with radius r, as in **Fig. 28.16a**. At every point on the circle, $\vec{B}$ and $d\vec{l}$ are parallel, and so $\vec{B} \cdot d\vec{l} = B\,dl$; since r is constant around the circle, B is constant as well. Alternatively, we can say that $B_\parallel$ is constant and equal to B at every point on the circle. Hence we can take B outside of the integral. The remaining integral is just the circumference of the circle, so

$$\oint \vec{B} \cdot d\vec{l} = \oint B_\parallel\, dl = B \oint dl = \frac{\mu_0 I}{2\pi r}(2\pi r) = \mu_0 I$$

The line integral is thus independent of the radius of the circle and is equal to μ_0 multiplied by the current passing through the area bounded by the circle.

In Fig. 28.16b the situation is the same, but the integration path now goes around the circle in the opposite direction. Now $\vec{B}$ and $d\vec{l}$ are antiparallel, so $\vec{B} \cdot d\vec{l} = -B\,dl$ and the line integral equals $-\mu_0 I$. We get the same result if the integration path is the same as in Fig. 28.16a, but the direction of the current is reversed. Thus $\oint \vec{B} \cdot d\vec{l}$ equals μ_0 multiplied by the current passing through the area bounded by the integration path, with a positive or negative sign depending on the direction of the current relative to the direction of integration.

There's a simple rule for the sign of the current; you won't be surprised to learn that it uses your right hand. Curl the fingers of your right hand around the integration path so that they curl in the direction of integration (that is, the direction that you use to evaluate $\oint \vec{B} \cdot d\vec{l}$). Then your right thumb indicates the positive current direction. Currents that pass through the integration path in this direction are positive; those in the opposite direction are negative. Using this rule, convince yourself that the current is positive in Fig. 28.16a and negative in Fig. 28.16b. Here's another way to say the same thing: Looking at the surface bounded by the integration path, integrate counterclockwise around the path as in Fig. 28.16a. Currents moving toward you through the surface are positive, and those going away from you are negative.

An integration path that does *not* enclose the conductor is used in Fig. 28.16c. Along the circular arc ab of radius r_1, $\vec{B}$ and $d\vec{l}$ are parallel and $B_\parallel = B_1 = \mu_0 I/2\pi r_1$; along the circular arc cd of radius r_2, $\vec{B}$ and $d\vec{l}$ are antiparallel and $B_\parallel = -B_2 = -\mu_0 I/2\pi r_2$. The $\vec{B}$ field is perpendicular to $d\vec{l}$ at each point on the straight sections bc and da, so $B_\parallel = 0$ and these sections contribute zero to the line integral. The total line integral is then

$$\oint \vec{B} \cdot d\vec{l} = \oint B_\parallel\, dl = B_1 \int_a^b dl + (0)\int_b^c dl + (-B_2)\int_c^d dl + (0)\int_d^a dl$$

$$= \frac{\mu_0 I}{2\pi r_1}(r_1\theta) + 0 - \frac{\mu_0 I}{2\pi r_2}(r_2\theta) + 0 = 0$$

The magnitude of $\vec{B}$ is greater on arc cd than on arc ab, but the arc length is less, so the contributions from the two arcs exactly cancel. Even though there is a magnetic field everywhere along the integration path, the line integral $\oint \vec{B} \cdot d\vec{l}$ is zero if there is no current passing through the area bounded by the path.

We can also derive these results for more general integration paths, such as the one in **Fig. 28.17a**. At the position of the line element $d\vec{l}$, the angle between $d\vec{l}$ and $\vec{B}$ is ϕ, and

$$\vec{B} \cdot d\vec{l} = B\,dl\cos\phi$$

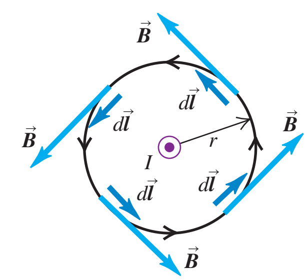
From the figure, $dl\cos\phi = r\,d\theta$, where $d\theta$ is the angle subtended by $d\vec{l}$ at the position of the conductor and r is the distance of $d\vec{l}$ from the conductor. Thus

$$\oint \vec{B} \cdot d\vec{l} = \oint \frac{\mu_0 I}{2\pi r}(r\,d\theta) = \frac{\mu_0 I}{2\pi}\oint d\theta$$

Figure **28.16** Three integration paths for the line integral of $\vec{B}$ in the vicinity of a long, straight conductor carrying current I *out* of the plane of the page (as indicated by the circle with a dot). The conductor is seen end-on.

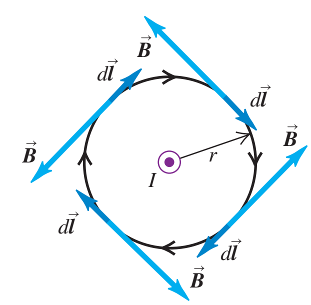
(a) Integration path is a circle centered on the conductor; integration goes around the circle counterclockwise.

Result: $\oint \vec{B} \cdot d\vec{l} = \mu_0 I$



(b) Same integration path as in (a), but integration goes around the circle clockwise.

Result: $\oint \vec{B} \cdot d\vec{l} = -\mu_0 I$



(c) An integration path that does not enclose the conductor

Result: $\oint \vec{B} \cdot d\vec{l} = 0$

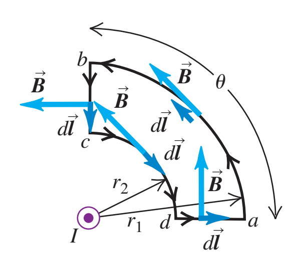
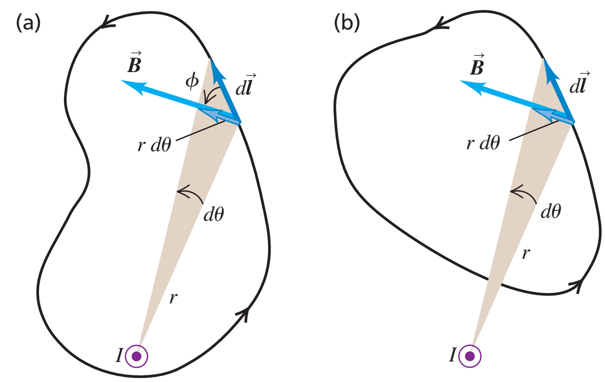


Figure 28.17 (a) A more general integration path for the line integral of $\vec{B}$ around a long, straight conductor carrying current I *out* of the plane of the page. The conductor is seen end-on. (b) A more general integration path that does not enclose the conductor.

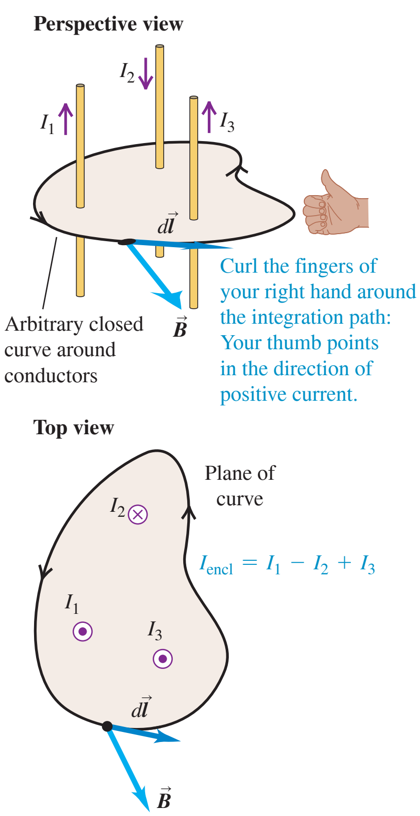


But $\oint d\theta$ is just equal to 2π, the total angle swept out by the radial line from the conductor to $d\vec{l}$ during a complete trip around the path. So we get

$$\oint \vec{B} \cdot d\vec{l} = \mu_0 I \tag{28.19}$$

This result doesn't depend on the shape of the path or on the position of the wire inside it. If the current in the wire is opposite to that shown, the integral has the opposite sign. But if the path doesn't enclose the wire (Fig. 28.17b), then the net change in θ during the trip around the integration path is zero; $\oint d\theta$ is zero instead of 2π and the line integral is zero.

Figure 28.18 Ampere's law.



Ampere's law: If we calculate the line integral of the magnetic field around a closed curve, the result equals μ_0 times the total enclosed current: $\oint \vec{B} \cdot d\vec{l} = \mu_0 I_{encl}$.

Ampere's Law: General Statement

We can generalize Ampere's law even further. Suppose *several* long, straight conductors pass through the surface bounded by the integration path. The total magnetic field $\vec{B}$ at any point on the path is the vector sum of the fields produced by the individual conductors. Thus the line integral of the total $\vec{B}$ equals μ_0 times the *algebraic sum* of the currents. In calculating this sum, we use the sign rule for currents described above. If the integration path does not enclose a particular wire, the line integral of the $\vec{B}$ field of that wire is zero, because the angle θ for that wire sweeps through a net change of zero rather than 2π during the integration. Any conductors present that are not enclosed by a particular path may still contribute to the value of $\vec{B}$ at every point, but the *line integrals* of their fields around the path are zero.

Thus we can replace I in Eq. (28.19) with I_{encl}, the algebraic sum of the currents *enclosed* or *linked* by the integration path, with the sum evaluated by using the sign rule just described (**Fig. 28.18**). Then **Ampere's law** says

Ampere's law:

$$\overbrace{\oint}^{\text{Line integral around a closed path}} \underbrace{\vec{B} \cdot d\vec{l}}_{\text{Scalar product of magnetic field and vector segment of path}} = \mu_0 I_{encl} \tag{28.20}$$

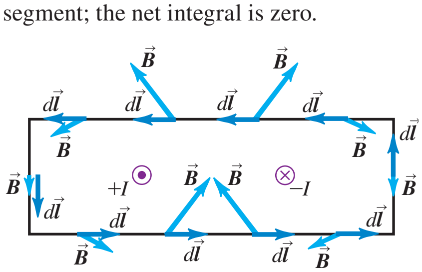
Magnetic constant (μ_0)
Net current enclosed by path (I_{encl})

While we have derived Ampere's law only for the special case of the field of several long, straight, parallel conductors, Eq. (28.20) is in fact valid for conductors and paths of *any* shape. The general derivation is no different in principle from what we have presented, but the geometry is more complicated.

If $\oint \vec{B} \cdot d\vec{l} = 0$, it *does not* necessarily mean that $\vec{B} = \mathbf{0}$ everywhere along the path, only that the total current through an area bounded by the path is zero. In Figs. 28.16c and 28.17b, the integration paths enclose no current at all; in **Fig. 28.19** there are positive and negative currents of equal magnitude through the area enclosed by the path. In both cases, $I_{encl} = 0$ and the line integral is zero.

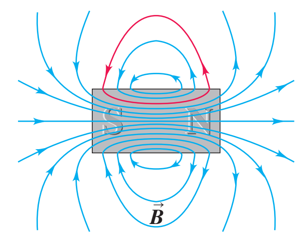
Equation (28.20) turns out to be valid *only* if the currents are steady and if no time-varying electric fields are present. In Chapter 29 we'll see how to generalize Ampere's law for time-varying fields.

Figure 28.19 An end-on view of two long, straight conductors carrying equal currents in opposite directions. The line integral $\oint \vec{B} \cdot d\vec{l}$ gets zero contribution from the upper and lower segments, a positive contribution from the left segment, and a negative contribution from the right segment; the net integral is zero.



CAUTION **Line integrals of electric and magnetic fields** In Chapter 23 we saw that the line integral of the electrostatic field $\vec{E}$ around any closed path is equal to zero; this is a statement that the electrostatic force $\vec{F} = q\vec{E}$ on a point charge q is conservative, so this force does zero work on a charge that moves around a closed path that returns to the starting point. The value of the line integral $\oint \vec{B} \cdot d\vec{l}$ is not similarly related to the question of whether the *magnetic* force is conservative. Remember that the magnetic force $\vec{F} = q\vec{v} \times \vec{B}$ on a moving charged particle is always *perpendicular* to $\vec{B}$, so $\oint \vec{B} \cdot d\vec{l}$ is *not* related to the work done by the magnetic force; as stated in Ampere's law, this integral is related only to the total current through a surface bounded by the integration path. In fact, the magnetic force on a moving charged particle is *not* conservative. A conservative force depends on only the position of the object on which the force is exerted, but the magnetic force on a moving charged particle also depends on the *velocity* of the particle.

TEST YOUR UNDERSTANDING OF SECTION 28.6 The accompanying figure shows magnetic field lines through the center of a permanent magnet. The magnet is not connected to a source of emf. One of the field lines is colored red. What can you conclude about the currents inside the permanent magnet within the region enclosed by this field line? (i) There are no currents inside the magnet; (ii) there are currents directed out of the plane of the page; (iii) there are currents directed into the plane of the page; (iv) not enough information is given to decide.



ANSWER

(ii) Imagine carrying out the integral $\oint \vec{B} \cdot d\vec{l}$ along an integration path that goes counterclockwise around the red magnetic field line. At each point along the path the magnetic field $\vec{B}$ and the infinitesimal segment $d\vec{l}$ are both tangent to the path, so $\vec{B} \cdot d\vec{l}$ is positive at each point and the integral $\oint \vec{B} \cdot d\vec{l}$ is likewise positive. It follows from Ampere's law $\oint \vec{B} \cdot d\vec{l} = \mu_0 I_{\text{encl}}$ and the right-hand rule that the integration path encloses a current directed out of the plane of the page. There are no currents in the empty space outside the magnet, so there must be currents inside the magnet (see Section 28.8).

28.7 APPLICATIONS OF AMPERE'S LAW

Below are several examples of using Ampere's law to calculate the magnetic field due to a current distribution. Problem-Solving Strategy 28.2 is directly analogous to Problem-Solving Strategy 22.1 (Section 22.4) for using Gauss's law to calculate the electric field due to a charge distribution; we suggest you review that strategy now.

CAUTION **Ampere's law is powerful but limited** Like Gauss's law for $\vec{E}$, Ampere's law for $\vec{B}$ can be used to calculate the field *only* in highly symmetric situations. As the next examples show, in such situations you can exploit the symmetry to calculate the line integral of $\vec{B}$ around a simple path.

PROBLEM-SOLVING STRATEGY 28.2 Ampere's Law

IDENTIFY *the relevant concepts:* Ampere's law, $\oint \vec{B} \cdot d\vec{l} = \mu_0 I_{\text{encl}}$, can yield the magnitude of $\vec{B}$ as a function of position if the geometry of the field-generating electric current is highly symmetric.

SET UP *the problem* using the following steps:

1. Determine the target variable(s). Usually one will be the magnitude of the $\vec{B}$ field as a function of position.
2. Select the integration path to use with Ampere's law. To determine the magnetic field at a certain point, the path must pass through that point. The integration path doesn't have to be any actual physical boundary; it may be in empty space, embedded in a solid object, or some of each. The integration path has to have enough *symmetry* to make evaluation of the integral possible. Ideally the path will be tangent to $\vec{B}$ in regions of interest; elsewhere the path should be perpendicular to $\vec{B}$ or should run through regions where $\vec{B} = \mathbf{0}$.

EXECUTE *the solution* as follows:

1. Carry out the integral $\oint \vec{B} \cdot d\vec{l}$ along the chosen path. If $\vec{B}$ is tangent to all or some portion of the path and has the same magnitude B at every point, then its line integral is the product of B and the length of that portion of the path. If $\vec{B}$ is perpendicular to some portion of the path, or if $\vec{B} = \mathbf{0}$, that portion makes no contribution to the integral.
2. In the integral $\oint \vec{B} \cdot d\vec{l}$, $\vec{B}$ is the *total* magnetic field at each point on the path; it can be caused by currents enclosed *or not enclosed* by the path. If *no* net current is enclosed by the path, the field at points on the path need not be zero, but the integral $\oint \vec{B} \cdot d\vec{l}$ is always zero.
3. Determine the current I_{encl} enclosed by the integration path. A right-hand rule gives the sign of this current: If you curl the fingers of your right hand so that they follow the path in the direction of integration, then your right thumb points in the direction of positive current. If $\vec{B}$ is tangent to the path everywhere and I_{encl} is positive, the direction of $\vec{B}$ is the same as the direction of integration. If instead I_{encl} is negative, $\vec{B}$ is in the direction opposite to that of the integration.
4. Use Ampere's law $\oint \vec{B} \cdot d\vec{l} = \mu_0 I$ to solve for the target variable.

EVALUATE *your answer:* If your result is an expression for the field magnitude as a function of position, check it by examining how the expression behaves in different limits.

EXAMPLE 28.7 Field of a long, straight, current-carrying conductor

In Section 28.6 we derived Ampere's law from Eq. (28.9) for the field $\vec{B}$ of a long, straight, current-carrying conductor. Reverse this process, and use Ampere's law to find $\vec{B}$ for this situation.

IDENTIFY and SET UP The situation has cylindrical symmetry, so in Ampere's law we take our integration path to be a circle with radius r centered on the conductor and lying in a plane perpendicular to it, as in Fig. 28.16a. The field $\vec{B}$ is everywhere tangent to this circle and has the same magnitude B everywhere on the circle.

EXECUTE With our choice of the integration path, Ampere's law [Eq. (28.20)] becomes

$$\oint \vec{B} \cdot d\vec{l} = \oint B_{\parallel}\, dl = B(2\pi r) = \mu_0 I$$

Equation (28.9), $B = \mu_0 I/2\pi r$, follows immediately.

Ampere's law determines the direction of $\vec{B}$ as well as its magnitude. Since we chose to go counterclockwise around the integration path, the positive direction for current is out of the plane of Fig. 28.16a; this is the same as the actual current direction in the figure, so I is positive and the integral $\oint \vec{B} \cdot d\vec{l}$ is also positive. Since the $d\vec{l}$'s run counterclockwise, the direction of $\vec{B}$ must be counterclockwise as well, as shown in Fig. 28.16a.

EVALUATE Our results are consistent with those in Section 28.6.

KEYCONCEPT Ampere's law is easiest to use if the magnetic field $\vec{B}$ is everywhere tangent to an integration path and has the same magnitude at all points along that path.

EXAMPLE 28.8 Field of a long cylindrical conductor

WITH VARIATION PROBLEMS

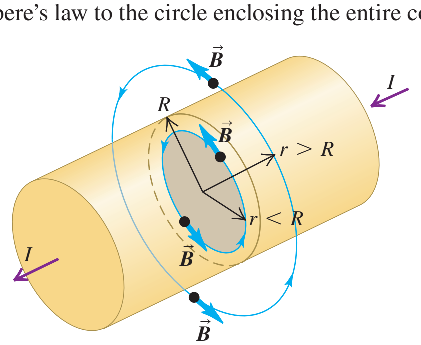
A cylindrical conductor with radius R carries a current I (**Fig. 28.20**). The current is uniformly distributed over the cross-sectional area of the conductor. Find the magnetic field as a function of the distance r from the conductor axis for points both inside $(r < R)$ and outside $(r > R)$ the conductor.

IDENTIFY and SET UP As in Example 28.7, the current distribution has cylindrical symmetry, and the magnetic field lines must be circles concentric with the conductor axis. To find the magnetic field inside and outside the conductor, we choose circular integration paths with radii $r < R$ and $r > R$, respectively (see Fig. 28.20).

EXECUTE In either case the field $\vec{B}$ has the same magnitude at every point on the circular integration path and is tangent to the path. Thus the magnitude of the line integral is simply $B(2\pi r)$. To find the current I_{encl} enclosed by a circular integration path inside the conductor $(r < R)$, note that the current density (current per unit area) is $J = I/\pi R^2$ so $I_{\text{encl}} = J(\pi r^2) = Ir^2/R^2$. Hence Ampere's law gives $B(2\pi r) = \mu_0 I r^2/R^2$, or

$$B = \frac{\mu_0 I}{2\pi}\frac{r}{R^2} \qquad \text{(inside the conductor, } r < R) \qquad (28.21)$$

Figure 28.20 To find the magnetic field at radius $r < R$, we apply Ampere's law to the circle enclosing the gray area. The current through the gray area is $(r^2/R^2)I$. To find the magnetic field at radius $r > R$, we apply Ampere's law to the circle enclosing the entire conductor.



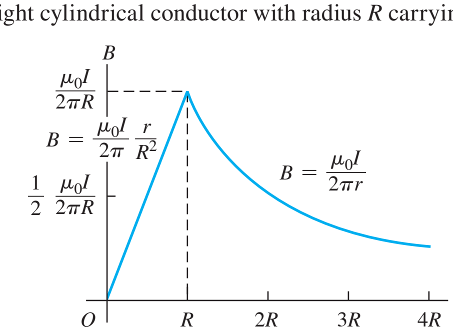
A circular integration path outside the conductor encloses the total current in the conductor, so $I_{\text{encl}} = I$. Applying Ampere's law gives the same equation as in Example 28.7, with the same result for B:

$$B = \frac{\mu_0 I}{2\pi r} \qquad \text{(outside the conductor, } r > R) \qquad (28.22)$$

Outside the conductor, the magnetic field is the same as that of a long, straight conductor carrying current I, independent of the radius R over which the current is distributed. Indeed, the magnetic field outside *any* cylindrically symmetric current distribution is the same as if the entire current were concentrated along the axis of the distribution. This is analogous to the results of Examples 22.5 and 22.9 (Section 22.4), in which we found that the *electric* field outside a spherically symmetric *charged* object is the same as though the entire charge were concentrated at the center.

EVALUATE At the surface of the conductor $(r = R)$, Eqs. (28.21) and (28.22) agree, as they must. **Figure 28.21** shows a graph of B as a function of r.

Figure 28.21 Magnitude of the magnetic field inside and outside a long, straight cylindrical conductor with radius R carrying a current I.



KEYCONCEPT The magnetic field lines produced by a solid, cylindrical, current-carrying conductor are concentric circles around the conductor, with the field direction given by a right-hand rule. Inside the conductor the field magnitude increases with increasing distance from the center; outside the conductor the magnitude decreases with distance.

EXAMPLE 28.9 Field of a solenoid

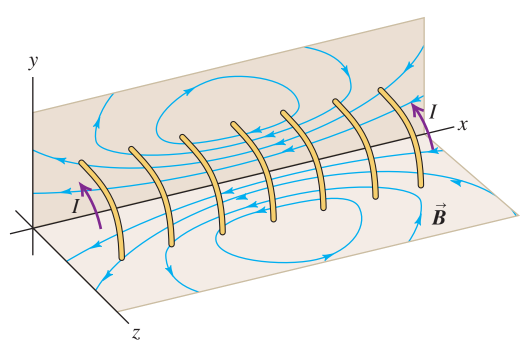
WITH VARIATION PROBLEMS

A **solenoid** consists of a helical winding of wire on a cylinder, usually circular in cross section. There can be thousands of closely spaced turns (often in several layers), each of which can be regarded as a circular loop. For simplicity, **Fig. 28.22** shows a solenoid with only a few turns. All turns carry the same current I, and the total $\vec{B}$ field at every point is the vector sum of the fields caused by the individual turns. The figure shows field lines in the xy- and xz-planes. We draw field lines that are uniformly spaced at the center of the solenoid. Exact calculations show that for a long, closely wound solenoid, half of these field lines emerge from the ends and half "leak out" through the windings between the center and the end, as the figure suggests.

If the solenoid is long in comparison with its cross-sectional diameter and the coils are tightly wound, the field inside the solenoid near its midpoint is very nearly uniform over the cross section and parallel to the axis; the *external* field near the midpoint is very small.

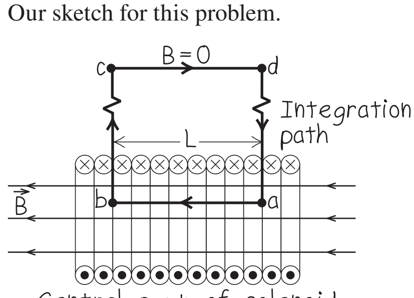
Use Ampere's law to find the field at or near the center of such a solenoid if it has n turns per unit length and carries current I.

Figure 28.22 Magnetic field lines produced by the current in a solenoid. For clarity, only a few turns are shown.



IDENTIFY and SET UP We assume that $\vec{B}$ is uniform inside the solenoid and zero outside. **Figure 28.23** shows the situation and our chosen integration path, rectangle $abcd$. Side ab, with length L, is parallel to the axis of the solenoid. Sides bc and da are taken to be very long so that side cd is far from the solenoid; then the field at side cd is negligibly small.

Figure 28.23 Our sketch for this problem.



EXECUTE Along side ab, $\vec{B}$ is parallel to the path and is constant. Our Ampere's-law integration takes us along side ab in the same direction as $\vec{B}$, so here $B_{\parallel} = +B$ and

$$\int_a^b \vec{B} \cdot d\vec{l} = BL$$

Along sides bc and da, $\vec{B}$ is perpendicular to the path, and so $B_{\parallel} = 0$; along side cd, $\vec{B} = \mathbf{0}$ and so $B_{\parallel} = 0$. Around the entire closed path, then, we have $\oint \vec{B} \cdot d\vec{l} = BL$.

In a length L there are nL turns, each of which passes once through $abcd$ carrying current I. Hence the total current enclosed by the rectangle is $I_{\text{encl}} = nLI$. The integral $\oint \vec{B} \cdot d\vec{l}$ is positive, so from Ampere's law I_{encl} must be positive as well. This means that the current passing through the surface bounded by the integration path must be in the direction shown in Fig. 28.23. Ampere's law then gives $BL = \mu_0 nLI$, or

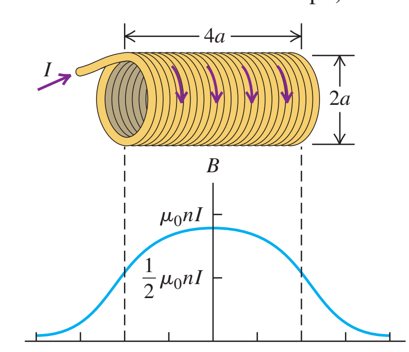
$$B = \mu_0 nI \qquad \text{(solenoid)} \qquad (28.23)$$

Side ab need not lie on the axis of the solenoid, so this result demonstrates that the field is uniform over the entire cross section at the center of the solenoid's length.

EVALUATE Note that the *direction* of $\vec{B}$ inside the solenoid is in the same direction as the solenoid's vector magnetic moment $\vec{\mu}$, as we found in Section 28.5 for a single current-carrying loop.

For points along the axis, the field is strongest at the center of the solenoid and drops off near the ends. For a solenoid very long in comparison to its diameter, the field magnitude at each end is exactly half that at the center. This is approximately the case even for a relatively short solenoid, as **Fig. 28.24** shows.

Figure 28.24 Magnitude of the magnetic field at points along the axis of a solenoid with length $4a$, equal to four times its radius a. The field magnitude at each end is about half its value at the center. (Compare with Fig. 28.14 for the field of N circular loops.)



KEYCONCEPT The magnetic field $\vec{B}$ of an ideal solenoid is uniform inside the solenoid and zero outside. The direction of $\vec{B}$ inside the solenoid is given by the same right-hand rule as for the field along the axis of a current-carrying coil (see Example 28.6).

EXAMPLE 28.10 Field of a toroidal solenoid

WITH VARIATION PROBLEMS

Figure 28.25a shows a doughnut-shaped **toroidal solenoid,** tightly wound with N turns of wire carrying a current I. (In a practical solenoid the turns would be much more closely spaced than they are in the figure.) Find the magnetic field at all points.

IDENTIFY and SET UP Ignoring the slight pitch of the helical windings, we can consider each turn of a tightly wound toroidal solenoid as a loop lying in a plane perpendicular to the large, circular axis of the toroid. The symmetry of the situation then tells us that the magnetic field lines must be circles concentric with the toroid axis. Therefore we choose circular integration paths for use with Ampere's law, so that the field $\vec{B}$ (if any) is tangent to each path at all points along the path. Figure 28.25b shows three such paths.

EXECUTE Along each path, $\oint \vec{B}\cdot d\vec{l}$ equals the product of B and the path circumference $l = 2\pi r$. The total current enclosed by path 1 is zero, so from Ampere's law the field $\vec{B} = \mathbf{0}$ everywhere on this path.

Each turn of the winding passes *twice* through the area bounded by path 3, carrying equal currents in opposite directions. The *net* current enclosed is therefore zero, and hence $\vec{B} = \mathbf{0}$ at all points on this path as well. We conclude that *the field of an ideal toroidal solenoid is confined to the space enclosed by the windings.* We can think of such a solenoid as a tightly wound, straight solenoid that has been bent into a circle.

For path 2, we have $\oint \vec{B}\cdot d\vec{l} = 2\pi rB$. Each turn of the winding passes *once* through the area bounded by this path, so $I_{encl} = NI$. We note that I_{encl} is positive for the clockwise direction of integration in Fig. 28.25b, so $\vec{B}$ is in the direction shown. Ampere's law then says that $2\pi rB = \mu_0 NI$, so

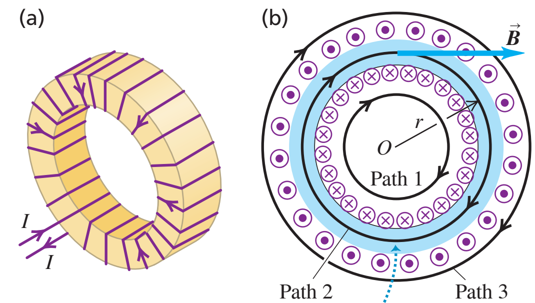
$$B = \frac{\mu_0 NI}{2\pi r} \quad \text{(toroidal solenoid)} \qquad (28.24)$$

EVALUATE Equation (28.24) indicates that B is *not* uniform over the interior of the core, because different points in the interior are different distances r from the toroid axis. However, if the radial extent of the core is small in comparison to r, the variation is slight. In that case, considering that $2\pi r$ is the circumferential length of the toroid and that $N/2\pi r$ is the number of turns per unit length n, the field may be written as $B = \mu_0 nI$, just as it is at the center of a long, *straight* solenoid.

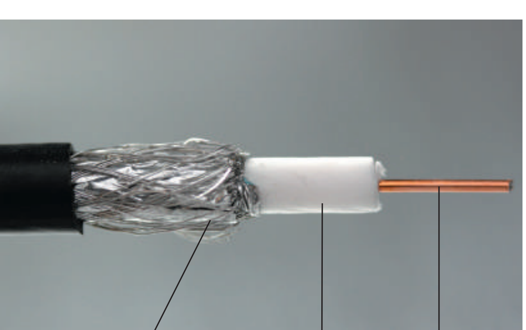
In a real toroidal solenoid the turns are not precisely circular loops but rather segments of a bent helix. As a result, the external field is not exactly zero. To estimate its magnitude, we imagine Fig. 28.25a as being *very* roughly equivalent, for points outside the torus, to a *single-turn* circular loop with radius r. At the center of such a loop, Eq. (28.17) gives $B = \mu_0 I/2r$; this is smaller than the field inside the solenoid by the factor N/π.

The equations we have derived for the field in a closely wound straight or toroidal solenoid are strictly correct only for windings in *vacuum*. For most practical purposes, however, they can be used for windings in air or on a core of any nonmagnetic, nonsuperconducting material. In the next section we'll show how these equations are modified if the core is a magnetic material.

Figure 28.25 (a) A toroidal solenoid. For clarity, only a few turns of the winding are shown. (b) Integration paths (black circles) used to compute the magnetic field $\vec{B}$ set up by the current (shown as dots and crosses).



KEYCONCEPT Inside the space enclosed by the windings of an ideal toroidal solenoid, the magnetic field lines are circles that follow the curvature of the toroid. Outside the windings, the magnetic field $\vec{B}$ is zero.



TEST YOUR UNDERSTANDING OF SECTION 28.7 In a *coaxial cable* (see photo), a conducting wire runs along the central axis of a hollow conducting cylinder. (The cable that connects a television set to a local cable provider is an example of a coaxial cable.) In such a cable a current I runs in one direction along the hollow conducting cylinder and is spread uniformly over the cylinder's cross-sectional area. An equal current runs in the opposite direction along the central wire. How does the magnitude B of the magnetic field outside such a cable depend on the distance r from the central axis of the cable? (i) B is proportional to $1/r$; (ii) B is proportional to $1/r^2$; (iii) B is zero at all points outside the cable.

ANSWER

(iii) By symmetry, any $\vec{B}$ field outside the cable must circulate around the cable, with circular field lines like those surrounding the solid cylindrical conductor in Fig. 28.20. Choose an integration path like the one shown in Fig. 28.20 with radius $r > R$, so that the path completely encloses the cable. As in Example 28.8, the integral $\oint \vec{B}\cdot d\vec{l}$ for this path has magnitude $B(2\pi r)$. From Ampere's law this is equal to $\mu_0 I_{encl}$. The net enclosed current I_{encl} is zero because it includes two currents of equal magnitude but opposite direction: one in the central wire and one in the hollow cylinder. Hence $B(2\pi r) = 0$, and so $B = 0$ for any value of r outside the cable. (The field is nonzero *inside* the cable; see Exercise 28.39.)

28.8 MAGNETIC MATERIALS

In discussing how currents cause magnetic fields, we have assumed that the conductors are surrounded by vacuum. But the coils in transformers, motors, generators, and electromagnets nearly always have iron cores to increase the magnetic field and confine it to desired regions. Permanent magnets, magnetic recording tapes, and computer disks depend directly on the magnetic properties of materials; when you store information on a computer disk, you are actually setting up an array of microscopic permanent magnets on the disk. So it is worthwhile to examine some aspects of the magnetic properties of materials. After describing the atomic origins of magnetic properties, we'll discuss three broad classes of magnetic behavior that occur in materials; these are called *paramagnetism, diamagnetism,* and *ferromagnetism.*

The Bohr Magneton

As we discussed briefly in Section 27.7, the atoms that make up all matter contain moving electrons, and these electrons form microscopic current loops that produce magnetic fields of their own. In many materials these currents are randomly oriented and cause no net magnetic field. But in some materials an external field (a field produced by currents outside the material) can cause these loops to become oriented preferentially with the field, so their magnetic fields *add* to the external field. We then say that the material is *magnetized.*

Let's look at how these microscopic currents come about. **Figure 28.26** shows a primitive model of an electron in an atom. We picture the electron (mass m, charge $-e$) as moving in a circular orbit with radius r and speed v. This moving charge is equivalent to a current loop. In Section 27.7 we found that a current loop with area A and current I has a magnetic dipole moment μ given by $\mu = IA$; for the orbiting electron the area of the loop is $A = \pi r^2$. To find the current associated with the electron, we note that the orbital period T (the time for the electron to make one complete orbit) is the orbit circumference divided by the electron speed: $T = 2\pi r/v$. The equivalent current I is the total charge passing any point on the orbit per unit time, which is just the magnitude e of the electron charge divided by the orbital period T:

$$I = \frac{e}{T} = \frac{ev}{2\pi r}$$

The magnetic moment $\mu = IA$ is then

$$\mu = \frac{ev}{2\pi r}(\pi r^2) = \frac{evr}{2} \tag{28.25}$$

It is useful to express μ in terms of the *angular momentum* L of the electron. For a particle moving in a circular path, the magnitude of angular momentum equals the magnitude of momentum mv multiplied by the radius r—that is, $L = mvr$ (see Section 10.5). Comparing this with Eq. (28.25), we can write

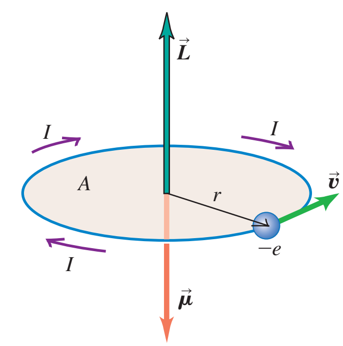
$$\mu = \frac{e}{2m}L \tag{28.26}$$

Equation (28.26) is useful in this discussion because atomic angular momentum is *quantized;* its component in a particular direction is always an integer multiple of $h/2\pi$, where h is a fundamental physical constant called *Planck's constant.* The numerical value of h is

$$h = 6.62607015 \times 10^{-34}\ \mathrm{J \cdot s}$$

The quantity $h/2\pi$ thus represents a fundamental unit of angular momentum in atomic systems, just as e is a fundamental unit of charge. Associated with the quantization of $\vec{L}$ is a fundamental uncertainty in the *direction* of $\vec{L}$ and therefore of $\vec{\mu}$. In the following discussion, when we speak of the magnitude of a magnetic moment, a more precise statement would be "maximum component in a given direction." Thus, to say that a magnetic moment $\vec{\mu}$ is aligned with a magnetic field $\vec{B}$ really means that $\vec{\mu}$ has its maximum possible component in the direction of $\vec{B}$; such components are always quantized.

Figure **28.26** An electron moving with speed v in a circular orbit of radius r has an angular momentum $\vec{L}$ and an oppositely directed orbital magnetic dipole moment $\vec{\mu}$. It also has a spin angular momentum and an oppositely directed spin magnetic dipole moment.



CAUTION Planck's constant and SI units We first introduced Planck's constant h in Section 1.3. Since h has units of $\mathrm{J \cdot s} = \mathrm{kg \cdot m^2/s}$, its *defined* value, combined with the definitions of the meter and the second, serves as the basis for the definition of the kilogram. Physicists define the kilogram this way because the value of h is relatively easy to measure in any physics laboratory, so it provides a useful standard for determining masses in kilograms.

Equation (28.26) shows that associated with the fundamental unit of angular momentum is a corresponding fundamental unit of magnetic moment. If $L = h/2\pi$, then

$$\mu = \frac{e}{2m}\left(\frac{h}{2\pi}\right) = \frac{eh}{4\pi m} \tag{28.27}$$

This quantity is called the **Bohr magneton,** denoted by μ_B. Its numerical value is

$$\mu_B = 9.274 \times 10^{-24}\ \text{A}\cdot\text{m}^2 = 9.274 \times 10^{-24}\ \text{J/T}$$

You should verify that these two sets of units are consistent. The second set is useful when we compute the potential energy $U = -\vec{\boldsymbol{\mu}} \cdot \vec{\boldsymbol{B}}$ for a magnetic moment in a magnetic field.

Electrons also have an intrinsic angular momentum, called *spin,* that is not related to orbital motion but that can be pictured in a classical model as spinning on an axis. This angular momentum also has an associated magnetic moment, and its magnitude turns out to be almost exactly one Bohr magneton. (Effects having to do with quantization of the electromagnetic field cause the spin magnetic moment to be about 1.001 μ_B.)

Paramagnetism

In an atom, most of the various orbital and spin magnetic moments of the electrons add up to zero. However, in some cases the atom has a net magnetic moment that is of the order of μ_B. When such a material is placed in a magnetic field, the field exerts a torque on each magnetic moment, as given by Eq. (27.26): $\vec{\boldsymbol{\tau}} = \vec{\boldsymbol{\mu}} \times \vec{\boldsymbol{B}}$. These torques tend to align the magnetic moments with the field, as we discussed in Section 27.7. In this position, the directions of the current loops are such as to *add* to the externally applied magnetic field.

We saw in Section 28.5 that the $\vec{\boldsymbol{B}}$ field produced by a current loop is proportional to the loop's magnetic dipole moment. In the same way, the additional $\vec{\boldsymbol{B}}$ field produced by microscopic electron current loops is proportional to the total magnetic moment $\vec{\boldsymbol{\mu}}_{\text{total}}$ per unit volume V in the material. We call this vector quantity the **magnetization** of the material, denoted by $\vec{\boldsymbol{M}}$:

$$\vec{\boldsymbol{M}} = \frac{\vec{\boldsymbol{\mu}}_{\text{total}}}{V} \tag{28.28}$$

The additional magnetic field due to magnetization of the material turns out to be equal simply to $\mu_0\vec{\boldsymbol{M}}$, where μ_0 is the same constant that appears in the law of Biot and Savart and Ampere's law. When such a material completely surrounds a current-carrying conductor, the total magnetic field $\vec{\boldsymbol{B}}$ in the material is

$$\vec{\boldsymbol{B}} = \vec{\boldsymbol{B}}_0 + \mu_0\vec{\boldsymbol{M}} \tag{28.29}$$

where $\vec{\boldsymbol{B}}_0$ is the field caused by the current in the conductor.

To check that the units in Eq. (28.29) are consistent, note that magnetization $\vec{\boldsymbol{M}}$ is the magnetic moment per unit volume. The units of magnetic moment are current times area $(\text{A}\cdot\text{m}^2)$, so the units of magnetization are $(\text{A}\cdot\text{m}^2)/\text{m}^3 = \text{A/m}$. From Section 28.1, the units of the constant μ_0 are $\text{T}\cdot\text{m/A}$. So the units of $\mu_0\vec{\boldsymbol{M}}$ are the same as the units of $\vec{\boldsymbol{B}}$: $(\text{T}\cdot\text{m/A})(\text{A/m}) = \text{T}$.

A material showing the behavior just described is said to be **paramagnetic.** The result is that the magnetic field at any point in such a material is greater by a dimensionless factor K_m, called the **relative permeability** of the material, than it would be if the material were replaced by vacuum. The value of K_m is different for different materials; for common paramagnetic solids and liquids at room temperature, K_m typically ranges from 1.00001 to 1.003.

All of the equations in this chapter that relate magnetic fields to their sources can be adapted to the situation in which the current-carrying conductor is embedded in a paramagnetic material. All that need be done is to replace μ_0 by $K_m\mu_0$. This product is usually denoted as μ and is called the **permeability** of the material:

$$\mu = K_m\mu_0 \tag{28.30}$$

CAUTION **Two meanings of the symbol μ** Equation (28.30) involves dangerous notation because we use μ for magnetic dipole moment as well as for permeability, as is customary. But beware: From now on, every time you see a μ, make sure you know whether it is permeability or magnetic moment. You can usually tell from the context.

The amount by which the relative permeability differs from unity is called the **magnetic susceptibility,** denoted by χ_m:

$$\chi_m = K_m - 1 \tag{28.31}$$

Both K_m and χ_m are dimensionless quantities. **Table 28.1** lists values of magnetic susceptibility for several materials. For example, for aluminum, $\chi_m = 2.2 \times 10^{-5}$ and $K_m = 1.000022$. The first group in the table consists of paramagnetic materials; we'll soon discuss the second group, which contains *diamagnetic* materials.

The tendency of atomic magnetic moments to align themselves parallel to the magnetic field (where the potential energy is minimum) is opposed by random thermal motion, which tends to randomize their orientations. For this reason, paramagnetic susceptibility always decreases with increasing temperature. In many cases it is inversely proportional to the absolute temperature T, and the magnetization M can be expressed as

$$M = C\frac{B}{T} \tag{28.32}$$

This relationship is called *Curie's law,* after its discoverer, Pierre Curie (1859–1906). The quantity C is a constant, different for different materials, called the *Curie constant.*

As we described in Section 27.7, an object with atomic magnetic dipoles is attracted to the poles of a magnet. In most paramagnetic substances this attraction is very weak due to thermal randomization of the atomic magnetic moments. But at very low temperatures the thermal effects are reduced, the magnetization increases in accordance with Curie's law, and the attractive forces are greater.

TABLE 28.1 Magnetic Susceptibilities of Paramagnetic and Diamagnetic Materials at $T = 20°C$

Material	$\chi_m = K_m - 1\ (\times 10^{-5})$
Paramagnetic	
Iron ammonium alum	66
Uranium	40
Platinum	26
Aluminum	2.2
Sodium	0.72
Oxygen gas	0.19
Diamagnetic	
Bismuth	−16.6
Mercury	−2.9
Silver	−2.6
Carbon (diamond)	−2.1
Lead	−1.8
Sodium chloride	−1.4
Copper	−1.0

EXAMPLE 28.11 Magnetic dipoles in a paramagnetic material

Nitric oxide (NO) is a paramagnetic compound. The magnetic moment of each NO molecule has a maximum component in any direction of about one Bohr magneton. Compare the interaction energy of such magnetic moments in a 1.5 T magnetic field with the average translational kinetic energy of molecules at 300 K.

IDENTIFY and SET UP This problem involves the energy of a magnetic moment in a magnetic field and the average thermal kinetic energy. We have Eq. (27.27), $U = -\vec{\boldsymbol{\mu}} \cdot \vec{\boldsymbol{B}}$, for the interaction energy of a magnetic moment $\vec{\boldsymbol{\mu}}$ with a $\vec{\boldsymbol{B}}$ field, and Eq. (18.16), $K = \frac{3}{2}kT$, for the average translational kinetic energy of a molecule at temperature T.

EXECUTE We can write $U = -\mu_{\parallel}B$, where $\mu_{\parallel}$ is the component of the magnetic moment $\vec{\boldsymbol{\mu}}$ in the direction of the $\vec{\boldsymbol{B}}$ field. Here the maximum value of $\mu_{\parallel}$ is about μ_B, so

$$\begin{aligned}|U|_{max} &\approx \mu_B B = (9.27 \times 10^{-24}\ \text{J/T})(1.5\ \text{T})\\ &= 1.4 \times 10^{-23}\ \text{J} = 8.7 \times 10^{-5}\ \text{eV}\end{aligned}$$

The average translational kinetic energy K is

$$\begin{aligned}K &= \tfrac{3}{2}kT = \tfrac{3}{2}(1.38 \times 10^{-23}\ \text{J/K})(300\ \text{K})\\ &= 6.2 \times 10^{-21}\ \text{J} = 0.039\ \text{eV}\end{aligned}$$

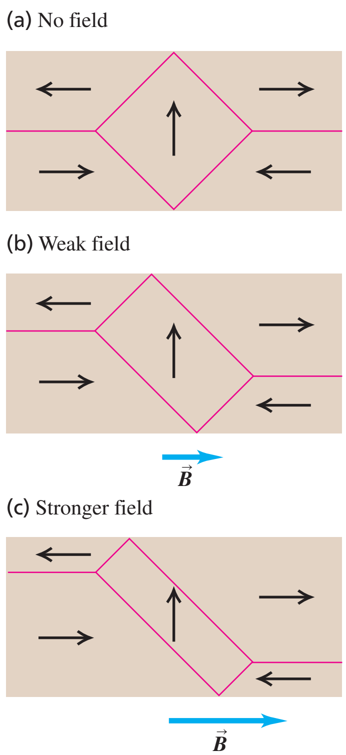
EVALUATE At 300 K the magnetic interaction energy is only about 0.2% of the thermal kinetic energy, so we expect only a slight degree of alignment. This is why paramagnetic susceptibilities at ordinary temperature are usually very small.

KEYCONCEPT In a paramagnetic material the magnetic dipole moments of its atoms tend to align with an externally applied magnetic field. Random thermal motions tend to cancel out this alignment, however. Paramagnetic materials are weakly attracted to the poles of a magnet.

Diamagnetism

In some materials the total magnetic moment of all the atomic current loops is zero when no magnetic field is present. But even these materials have magnetic effects because an external field alters electron motions within the atoms, causing additional current loops and induced magnetic dipoles comparable to the induced *electric* dipoles we studied in Section 28.5. In this case the additional field caused by these current loops is always *opposite* in direction to that of the external field. (This behavior is explained by Faraday's law of induction, which we'll study in Chapter 29. An induced current always tends to cancel the field change that caused it.)

Figure 28.27 In this drawing adapted from a magnified photo, the arrows show the directions of magnetization in the domains of a single crystal of nickel. Domains that are magnetized in the direction of an applied magnetic field grow larger.



Such materials are said to be **diamagnetic.** They always have negative susceptibility, as shown in Table 28.1, and relative permeability K_m slightly *less* than unity, typically of the order of 0.99990 to 0.99999 for solids and liquids. Diamagnetic susceptibilities are very nearly temperature independent.

Ferromagnetism

There is a third class of materials, called **ferromagnetic** materials, that includes iron, nickel, cobalt, and many alloys containing these elements. In these materials, strong interactions between atomic magnetic moments cause them to line up parallel to each other in regions called **magnetic domains,** even when no external field is present. **Figure 28.27** shows an example of magnetic domain structure. Within each domain, nearly all of the atomic magnetic moments are parallel.

When there is no externally applied field, the domain magnetizations are randomly oriented. But when a field $\vec{B}_0$ (caused by external currents) is present, the domains tend to orient themselves parallel to the field. The domain boundaries also shift; the domains that are magnetized in the field direction grow, and those that are magnetized in other directions shrink. Because the total magnetic moment of a domain may be many thousands of Bohr magnetons, the torques that tend to align the domains with an external field are much stronger than occur with paramagnetic materials. The relative permeability K_m is *much* larger than unity, typically of the order of 1000 to 100,000. As a result, an object made of a ferromagnetic material such as iron is strongly magnetized by the field from a permanent magnet and is attracted to the magnet (see Fig. 27.38). A paramagnetic material such as aluminum is also attracted to a permanent magnet, but K_m for paramagnetic materials is so much smaller for such a material than for ferromagnetic materials that the attraction is very weak. Thus a magnet can pick up iron nails, but not aluminum cans.

As the external field is increased, a point is eventually reached at which nearly *all* the magnetic moments in the ferromagnetic material are aligned parallel to the external field. This condition is called *saturation magnetization;* after it is reached, further increase in the external field causes no increase in magnetization or in the additional field caused by the magnetization.

Figure 28.28 shows a "magnetization curve," a graph of magnetization M as a function of external magnetic field B_0, for soft iron. An alternative description of this behavior is that K_m is not constant but decreases as B_0 increases. (Paramagnetic materials also show saturation at sufficiently strong fields. But the magnetic fields required are so large that departures from a linear relationship between M and B_0 in these materials can be observed only at very low temperatures, 1 K or so.)

For many ferromagnetic materials the relationship of magnetization to external magnetic field is different when the external field is increasing from when it is decreasing. **Figure 28.29a** shows this relationship for such a material. When the material is magnetized to saturation and then the external field is reduced to zero, some magnetization

Figure 28.28 A magnetization curve for a ferromagnetic material. The magnetization M approaches its saturation value M_{sat} as the magnetic field B_0 (caused by external currents) becomes large.

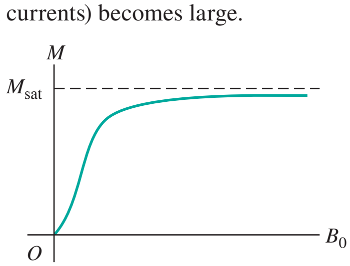
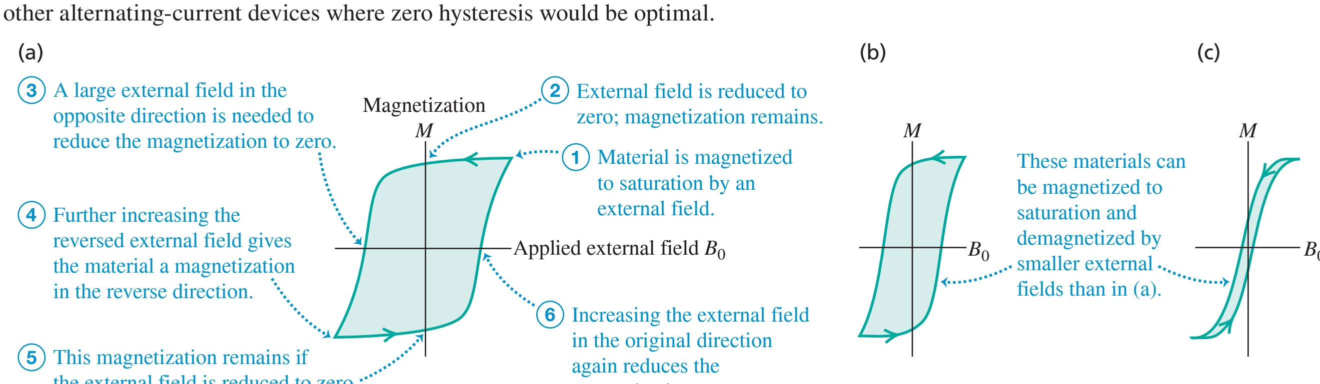


Figure 28.29 Hysteresis loops. The materials of both (a) and (b) remain strongly magnetized when B_0 is reduced to zero. Since (a) is also hard to demagnetize, it would be good for permanent magnets. Since (b) magnetizes and demagnetizes more easily, it could be used as a computer memory material. The material of (c) would be useful for transformers and other alternating-current devices where zero hysteresis would be optimal.

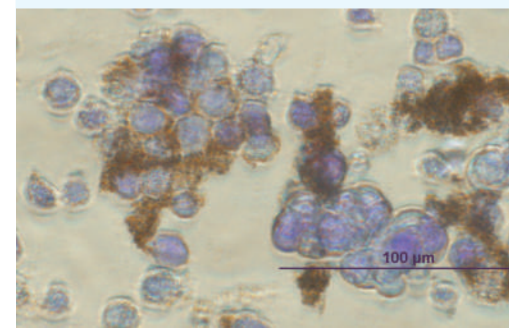


remains. This behavior is characteristic of permanent magnets, which retain most of their saturation magnetization when the magnetizing field is removed. To reduce the magnetization to zero requires a magnetic field in the reverse direction.

This behavior is called **hysteresis,** and the curves in Fig. 28.29 are called *hysteresis loops.* Magnetizing and demagnetizing a material that has hysteresis involve the dissipation of energy, and the temperature of the material increases during such a process.

Ferromagnetic materials are widely used in electromagnets, transformer cores, and motors and generators, in which it is desirable to have as large a magnetic field as possible for a given current. Because hysteresis dissipates energy, materials that are used in these applications should usually have as narrow a hysteresis loop as possible. Soft iron is often used; it has high permeability without appreciable hysteresis. For permanent magnets a broad hysteresis loop is usually desirable, with large zero-field magnetization and large reverse field needed to demagnetize. Many kinds of steel and many alloys, such as Alnico, are used for permanent magnets. The remaining magnetic field in such a material, after it has been magnetized to near saturation, is typically of the order of 1 T, corresponding to a remaining magnetization $M = B/\mu_0$ of about 800,000 A/m.

BIO APPLICATION Ferromagnetic Nanoparticles for Cancer Therapy The violet blobs in this microscope image are cancer cells that have broken away from a tumor and threaten to spread throughout a patient's body. An experimental technique for fighting these cells uses particles of a ferromagnetic material (shown in brown) injected into the body. These particles are coated with a chemical that preferentially attaches to cancer cells. A magnet outside the patient then "steers" the particles out of the body, taking the cancer cells with them. (Photo courtesy of cancer researcher Dr. Kenneth Scarberry.)



EXAMPLE 28.12 A ferromagnetic material

A cube-shaped permanent magnet is made of a ferromagnetic material with a magnetization M of about 8×10^5 A/m. The side length is 2 cm. (a) Find the magnetic dipole moment of the magnet. (b) Estimate the magnetic field due to the magnet at a point 10 cm from the magnet along its axis.

IDENTIFY and SET UP This problem uses the relationship between magnetization M and magnetic dipole moment μ_{total} and the idea that a magnetic dipole produces a magnetic field. We find μ_{total} from Eq. (28.28). To estimate the field, we approximate the magnet as a current loop with this same magnetic moment and use Eq. (28.18).

EXECUTE (a) From Eq. (28.28),

$$\mu_{\text{total}} = MV = (8 \times 10^5\ \text{A/m})(2 \times 10^{-2}\ \text{m})^3 = 6\ \text{A} \cdot \text{m}^2$$

(b) From Eq. (28.18), the magnetic field on the axis of a current loop with magnetic moment μ_{total} is

$$B = \frac{\mu_0 \mu_{\text{total}}}{2\pi (x^2 + a^2)^{3/2}}$$

where x is the distance from the loop and a is its radius. We can use this expression here if we take a to refer to the size of the permanent magnet. Strictly speaking, there are complications because our magnet does not have the same geometry as a circular current loop. But because $x = 10$ cm is fairly large in comparison to the 2 cm size of the magnet, the term a^2 is negligible in comparison to x^2 and can be ignored. So

$$B \approx \frac{\mu_0 \mu_{\text{total}}}{2\pi x^3} = \frac{(4\pi \times 10^{-7}\ \text{T} \cdot \text{m/A})(6\ \text{A} \cdot \text{m}^2)}{2\pi (0.1\ \text{m})^3} = 1 \times 10^{-3}\ \text{T} = 10\ \text{G}$$

which is about ten times stronger than the earth's magnetic field.

EVALUATE We calculated B at a point *outside* the magnetic material and therefore used μ_0, not the permeability μ of the magnetic material, in our calculation. You would substitute permeability μ for μ_0 if you were calculating B *inside* a material with relative permeability K_m, for which $\mu = K_m \mu_0$.

KEYCONCEPT In a ferromagnetic material there are strong interactions between the magnetic dipole moments of neighboring atoms, and the material can maintain a large overall magnetic dipole even if there is no externally applied magnetic field.

TEST YOUR UNDERSTANDING OF SECTION 28.8 Which of the following materials are attracted to a magnet? (i) Sodium; (ii) bismuth; (iii) lead; (iv) uranium.

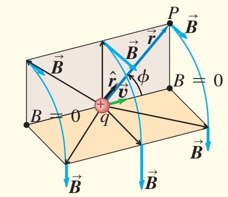
ANSWER

(i), (iv) Sodium and uranium are paramagnetic materials and hence are attracted to a magnet, while bismuth and lead are diamagnetic materials that are repelled by a magnet. (See Table 28.1.)

CHAPTER 28 SUMMARY

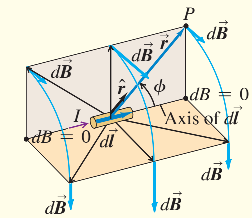
Magnetic field of a moving charge: The magnetic field $\vec{B}$ created by a charge q moving with velocity $\vec{v}$ depends on the distance r from the source point (the location of q) to the field point (where $\vec{B}$ is measured). The $\vec{B}$ field is perpendicular to $\vec{v}$ and to $\hat{r}$, the unit vector directed from the source point to the field point. The principle of superposition of magnetic fields states that the total $\vec{B}$ field produced by several moving charges is the vector sum of the fields produced by the individual charges. (See Example 28.1.)

$$\vec{B} = \frac{\mu_0}{4\pi}\frac{q\vec{v} \times \hat{r}}{r^2} \tag{28.2}$$



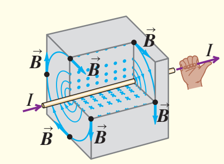
Magnetic field of a current-carrying conductor: The law of Biot and Savart gives the magnetic field $d\vec{B}$ created by an element $d\vec{l}$ of a conductor carrying current I. The field $d\vec{B}$ is perpendicular to both $d\vec{l}$ and $\hat{r}$, the unit vector from the element to the field point. The $\vec{B}$ field created by a finite current-carrying conductor is the integral of $d\vec{B}$ over the length of the conductor. (See Example 28.2.)

$$d\vec{B} = \frac{\mu_0}{4\pi}\frac{I\,d\vec{l} \times \hat{r}}{r^2} \tag{28.6}$$



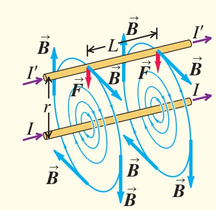
Magnetic field of a long, straight, current-carrying conductor: The magnetic field $\vec{B}$ at a distance r from a long, straight conductor carrying a current I has a magnitude that is inversely proportional to r. The magnetic field lines are circles coaxial with the wire, with directions given by the right-hand rule. (See Examples 28.3 and 28.4.)

$$B = \frac{\mu_0 I}{2\pi r} \tag{28.9}$$



Magnetic force between current-carrying conductors: Two long, parallel, current-carrying conductors attract if the currents are in the same direction and repel if the currents are in opposite directions. The magnetic force per unit length between the conductors depends on their currents I and I' and separation r. The definition of the ampere is based on this relationship. (See Example 28.5.)

$$\frac{F}{L} = \frac{\mu_0 I I'}{2\pi r} \tag{28.11}$$



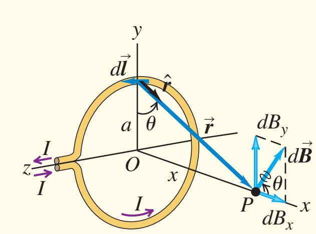
Magnetic field of a current loop: The law of Biot and Savart allows us to calculate the magnetic field produced along the axis of a circular conducting loop of radius a carrying current I. The field depends on the distance x along the axis from the center of the loop to the field point. If there are N loops, the field is multiplied by N. At the center of the loop, $x = 0$. (See Example 28.6.)

$$B_x = \frac{\mu_0 I a^2}{2(x^2 + a^2)^{3/2}} \tag{28.15}$$

(circular loop)

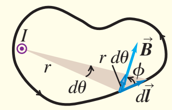
$$B_x = \frac{\mu_0 N I}{2a} \tag{28.17}$$

(center of N circular loops)



Ampere's law: Ampere's law states that the line integral of $\vec{B}$ around any closed path equals μ_0 times the net current through the area enclosed by the path. The positive sense of current is determined by a right-hand rule. (See Examples 28.7–28.10.)

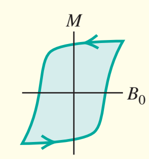
$$\oint \vec{B} \cdot d\vec{l} = \mu_0 I_{\text{encl}} \tag{28.20}$$



Magnetic fields due to current distributions: The table lists magnetic fields caused by several current distributions. In each case the conductor is carrying current I.

Current Distribution	Point in Magnetic Field	Magnetic-Field Magnitude
Long, straight conductor	Distance r from conductor	$B = \dfrac{\mu_0 I}{2\pi r}$
Circular loop of radius a (for N loops, multiply these expressions for magnetic-field amplitude by N)	On axis of loop	$B = \dfrac{\mu_0 I a^2}{2(x^2 + a^2)^{3/2}}$
	At center of loop	$B = \dfrac{\mu_0 I}{2a}$
Long cylindrical conductor of radius R	Inside conductor, $r < R$	$B = \dfrac{\mu_0 I}{2\pi} \dfrac{r}{R^2}$
	Outside conductor, $r > R$	$B = \dfrac{\mu_0 I}{2\pi r}$
Long, closely wound solenoid with n turns per unit length, near its midpoint	Inside solenoid, near center	$B = \mu_0 n I$
	Outside solenoid	$B \approx 0$
Tightly wound toroidal solenoid (toroid) with N turns	Within the space enclosed by the windings, distance r from symmetry axis	$B = \dfrac{\mu_0 N I}{2\pi r}$
	Outside the space enclosed by the windings	$B \approx 0$

Magnetic materials: When magnetic materials are present, the magnetization of the material causes an additional contribution to $\vec{B}$. For paramagnetic and diamagnetic materials, μ_0 is replaced in magnetic-field expressions by $\mu = K_m \mu_0$, where μ is the permeability of the material and K_m is its relative permeability. The magnetic susceptibility χ_m is defined as $\chi_m = K_m - 1$. Magnetic susceptibilities for paramagnetic materials are small positive quantities; those for diamagnetic materials are small negative quantities. For ferromagnetic materials, K_m is much larger than unity and is not constant. Some ferromagnetic materials are permanent magnets, retaining their magnetization even after the external magnetic field is removed. (See Examples 28.11 and 28.12.)



Chapter 28 Media Assets

GUIDED PRACTICE

For assigned homework and other learning materials, go to Mastering Physics.

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLE 28.1 (Section 28.1) and EXAMPLE 28.2 (Section 28.2) before attempting these problems.

VP28.2.1 A proton with charge $+1.60 \times 10^{-19}$ C is at the origin ($x = y = z = 0$) and is moving at 2.00×10^5 m/s in the $+x$-direction. Find the magnetic field $\vec{B}$ that this proton produces at (a) $x = 0$, $y = 1.00$ mm, $z = 0$; (b) $x = 0$, $y = 0$, $z = 2.00$ mm; (c) $x = 1.00$ mm, $y = 0$, $z = 0$; (d) $x = 1.00$ mm, $y = 1.00$ mm, $z = 0$.

VP28.2.2 A proton (charge $+1.60 \times 10^{-19}$ C) and an electron (charge -1.60×10^{-19} C) are both moving in the xy-plane with the same speed, 4.20×10^5 m/s. The proton is moving in the $+y$-direction along the line $x = 0$, and the electron is moving in the $-y$-direction along the line $x = +2.00$ mm. At the instant when the proton and electron are at their closest approach, what are the magnitude and direction of the magnetic force that (a) the proton exerts on the electron and (b) the electron exerts on the proton?

VP28.2.3 A segment of wire centered at the origin ($x = y = z = 0$) is 2.00 mm in length and carries a current of 6.00 A in the $+y$-direction. You measure the magnetic field due to this segment at the point $x = 3.00$ m, $y = 4.00$ m, $z = 0$. Find (a) the unit vector from the wire segment to this point and (b) the magnetic field at this point.

VP28.2.4 A segment of wire centered at the origin ($x = y = z = 0$) is 8.00 mm in length and carries a current of 4.00 A in the $+z$-direction. Find the magnetic force that this segment exerts on an electron (charge -1.60×10^{-19} C) at the point $x = 1.25$ m, $y = 0$, $z = 0$ that is moving at 3.00×10^5 m/s in the $-x$-direction.

Be sure to review EXAMPLES 28.3 and 28.4 (Section 28.3) and EXAMPLE 28.5 (Section 28.4) before attempting these problems.

VP28.5.1 How much current must a long, straight wire carry in order for the magnetic field that it produces 1.40 cm from the central axis of the wire to have magnitude 3.20×10^{-5} T?

VP28.5.2 Two long, straight, conducting wires are both perpendicular to the xy-plane. Wire 1 passes through the point $x = 0$, $y = 1.00$ cm, $z = 0$ and carries current 1.00 A in the $+z$-direction. Wire 2 passes through the point $x = 0$, $y = -1.00$ cm, $z = 0$ and carries current 4.00 A also in the $+z$-direction. Find the net magnetic field due to both wires at (a) $x = 0$, $y = 0$, $z = 0$; (b) $x = 0$, $y = 2.00$ cm, $z = 0$; (c) $x = 0$, $y = -2.00$ cm, $z = 0$.

VP28.5.3 A long, straight, conducting wire carries a current of 2.00 A. A second, identical wire is parallel to the first and separated from it by 3.00 cm. The magnetic field at a point halfway between the two wires has magnitude 4.00×10^{-5} T. What is the current in the second wire if its direction is (a) the same as in the first wire? (b) Opposite to that in the first wire?

VP28.5.4 Two straight, parallel, superconducting wires are 0.900 m apart. Both wires are oriented straight up and down. One wire carries a current of 1.30×10^4 A upward and experiences a magnetic force per unit length of 95.0 N/m that attracts it to the other wire. What are the magnitude and direction of the current in the other wire?

Be sure to review EXAMPLES 28.8, 28.9, and 28.10 (Section 28.7) before attempting these problems.

VP28.10.1 A solid cylindrical conductor with radius 4.50 cm carries a current of 2.00 A. The current is uniformly distributed over the cross-sectional area of the conductor. Find the magnitude of the magnetic field at the following distances from the conductor axis: (a) 2.50 cm; (b) 4.50 cm; (c) 6.00 cm.

VP28.10.2 A long, hollow cylinder with inner radius R_1 and outer radius R_2 carries current along its length. The current is uniformly distributed over the cross-sectional area of the cylinder and has current density J. (a) Find the magnetic-field magnitude B as a function of the distance r from the conductor axis for points inside the hollow interior ($r < R_1$), inside the solid conductor ($R_2 > r > R_1$), and outside the conductor ($r > R_2$). (b) For which value of r is B greatest?

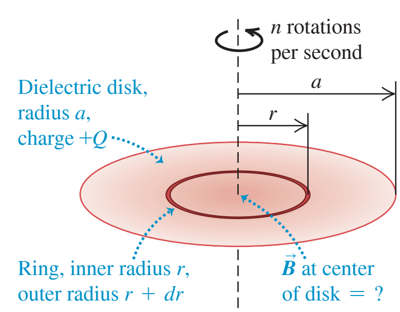
VP28.10.3 A very long solenoid with tightly wound coils has 8.00×10^3 turns of wire per meter of length. What must the current be in the solenoid in order for the field inside the solenoid to have magnitude 0.0320 T?

VP28.10.4 A copper wire carries a current of 0.840 A. You wrap this wire around a hollow toroidal core, forming a toroidal solenoid. The inner and outer radii of the core are 6.00 cm and 8.00 cm, respectively. (a) How many turns of wire around the core are required to produce a magnetic field of magnitude 2.00×10^{-3} T at a distance of 7.00 cm from the toroid axis? (b) With this number of turns, what are the maximum and minimum field magnitudes in the interior of the core?

BRIDGING PROBLEM Magnetic Field of a Charged, Rotating Dielectric Disk

A thin dielectric disk with radius a has a total charge $+Q$ distributed uniformly over its surface (**Fig. 28.30**). It rotates n times per second about an axis perpendicular to the surface of the disk and passing through its center. Find the magnetic field at the center of the disk.

Figure 28.30 Finding the $\vec{B}$ field at the center of a uniformly charged, rotating disk.



SOLUTION GUIDE

IDENTIFY and SET UP

1. Think of the rotating disk as a series of concentric rotating rings. Each ring acts as a circular current loop that produces a magnetic field at the center of the disk.
2. Use the results of Section 28.5 to find the magnetic field due to a single ring. Then integrate over all rings to find the total field.

EXECUTE

3. Find the charge on a ring with inner radius r and outer radius $r + dr$ (Fig. 28.30).
4. How long does it take the charge found in step 3 to make a complete trip around the rotating ring? Use this to find the current of the rotating ring.
5. Use a result from Section 28.5 to determine the magnetic field that this ring produces at the center of the disk.
6. Integrate your result from step 5 to find the total magnetic field from all rings with radii from $r = 0$ to $r = a$.

EVALUATE

7. Does your answer have the correct units?
8. Suppose all of the charge were concentrated at the rim of the disk (at $r = a$). Would this increase or decrease the field at the center of the disk?

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q28.1 A topic of current interest in physics research is the search (thus far unsuccessful) for an isolated magnetic pole, or magnetic *monopole*. If such an entity were found, how could it be recognized? What would its properties be?

Q28.2 Streams of charged particles emitted from the sun during periods of solar activity create a disturbance in the earth's magnetic field. How does this happen?

Q28.3 The text discussed the magnetic field of an infinitely long, straight conductor carrying a current. Of course, there is no such thing as an infinitely long *anything*. How do you decide whether a particular wire is long enough to be considered infinite?

Q28.4 Two parallel conductors carrying current in the same direction attract each other. If they are permitted to move toward each other, the forces of attraction do work. From where does the energy come? Does this contradict the assertion in Chapter 27 that magnetic forces on moving charges do no work? Explain.

Q28.5 Pairs of conductors carrying current into or out of the power-supply components of electronic equipment are sometimes twisted together to reduce magnetic-field effects. Why does this help?

Q28.6 Suppose you have three long, parallel wires arranged so that in cross section they are at the corners of an equilateral triangle. Is there any way to arrange the currents so that all three wires attract each other? So that all three wires repel each other? Explain.

Q28.7 In deriving the force on one of the long, current-carrying conductors in Section 28.4, why did we use the magnetic field due to only one of the conductors? That is, why didn't we use the *total* magnetic field due to *both* conductors?

Q28.8 Two concentric, coplanar, circular loops of wire of different diameter carry currents in the same direction. Describe the nature of the force exerted on the inner loop by the outer loop and on the outer loop by the inner loop.

Q28.9 A current was sent through a helical coil spring. The spring contracted, as though it had been compressed. Why?

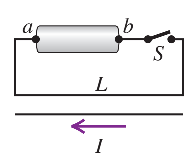
Q28.10 What are the relative advantages and disadvantages of Ampere's law and the law of Biot and Savart for practical calculations of magnetic fields?

Q28.11 Magnetic field lines never have a beginning or an end. Use this to explain why it is reasonable for the field of an ideal toroidal solenoid to be confined entirely to its interior, while a straight solenoid *must* have some field outside.

Q28.12 Two very long, parallel wires carry equal currents in opposite directions. (a) Is there any place that their magnetic fields completely cancel? If so, where? If not, why not? (b) How would the answer to part (a) change if the currents were in the same direction?

Q28.13 In the circuit shown in **Fig. Q28.13**, when switch S is suddenly closed, the wire L is pulled toward the lower wire carrying current I. Which (a or b) is the positive terminal of the battery? How do you know?

Figure **Q28.13**



Q28.14 A metal ring carries a current that causes a magnetic field B_0 at the center of the ring and a field B at point P a distance x from the center along the axis of the ring. If the radius of the ring is doubled, find the magnetic field at the center. Will the field at point P change by the same factor? Why?

Q28.15 Show that the units $\mathrm{A \cdot m^2}$ and J/T for the Bohr magneton are equivalent.

Q28.16 Why should the permeability of a paramagnetic material be expected to decrease with increasing temperature?

Q28.17 If a magnet is suspended over a container of liquid air, it attracts droplets to its poles. The droplets contain only liquid oxygen; even though nitrogen is the primary constituent of air, it is not attracted to the magnet. Explain what this tells you about the magnetic susceptibilities of oxygen and nitrogen, and explain why a magnet in ordinary, room-temperature air doesn't attract molecules of oxygen *gas* to its poles.

Q28.18 What features of atomic structure determine whether an element is diamagnetic or paramagnetic? Explain.

Q28.19 The magnetic susceptibility of paramagnetic materials is quite strongly temperature dependent, but that of diamagnetic materials is nearly independent of temperature. Why the difference?

Q28.20 A cylinder of iron is placed so that it is free to rotate around its axis. Initially the cylinder is at rest, and a magnetic field is applied to the cylinder so that it is magnetized in a direction parallel to its axis. If the direction of the *external* field is suddenly reversed, the direction of magnetization will also reverse and the cylinder will begin rotating around its axis. (This is called the *Einstein–de Haas effect*.) Explain why the cylinder begins to rotate.

EXERCISES

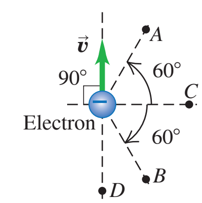
Section 28.1 Magnetic Field of a Moving Charge

28.1 •• A $+6.00\ \mu\mathrm{C}$ point charge is moving at a constant 8.00×10^6 m/s in the $+y$-direction, relative to a reference frame. At the instant when the point charge is at the origin of this reference frame, what is the magnetic-field vector $\vec{B}$ it produces at the following points: (a) $x = 0.500$ m, $y = 0$, $z = 0$; (b) $x = 0$, $y = -0.500$ m, $z = 0$; (c) $x = 0$, $y = 0$, $z = +0.500$ m; (d) $x = 0$, $y = -0.500$ m, $z = +0.500$ m?

28.2 • **Fields Within the Atom.** In the Bohr model of the hydrogen atom, the electron moves in a circular orbit of radius 5.3×10^{-11} m with a speed of 2.2×10^6 m/s. If we are viewing the atom in such a way that the electron's orbit is in the plane of the paper with the electron moving clockwise, find the magnitude and direction of the electric and magnetic fields that the electron produces at the location of the nucleus (treated as a point).

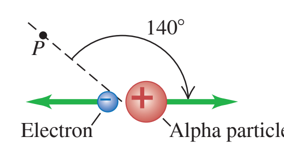
28.3 • An electron moves at $0.100c$ as shown in **Fig. E28.3**. Find the magnitude and direction of the magnetic field this electron produces at the following points, each 2.00 μm from the electron: (a) points A and B; (b) point C; (c) point D.

Figure **E28.3**



28.4 •• An alpha particle (charge $+2e$) and an electron move in opposite directions from the same point, each with the speed of 2.50×10^5 m/s (**Fig. E28.4**). Find the magnitude and direction of the total magnetic field these charges produce at point P, which is 8.65 nm from each charge.

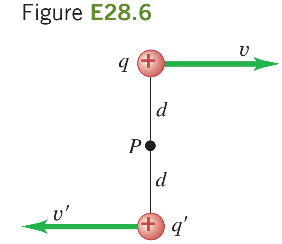
Figure **E28.4**



28.5 • A $-4.80\ \mu\mathrm{C}$ charge is moving at a constant speed of 6.80×10^5 m/s in the $+x$-direction relative to a reference frame. At the instant when the point charge is at the origin, what is the magnetic-field vector it produces at the following points: (a) $x = 0.500$ m, $y = 0$, $z = 0$; (b) $x = 0$, $y = 0.500$ m, $z = 0$; (c) $x = 0.500$ m, $y = 0.500$ m, $z = 0$; (d) $x = 0$, $y = 0$, $z = 0.500$ m?

28.6 • Positive point charges $q = +8.00\ \mu\mathrm{C}$ and $q' = +3.00\ \mu\mathrm{C}$ are moving relative to an observer at point P, as shown in **Fig. E28.6**. The distance d is 0.120 m, $v = 4.50 \times 10^6$ m/s, and $v' = 9.00 \times 10^6$ m/s. (a) When the two charges are at the locations shown in the figure, what are the magnitude and direction of the net magnetic field they produce at point P? (b) What are the magnitude and direction of the electric and magnetic forces that each charge exerts on the other, and what is the ratio of the magnitude of the electric force to the magnitude of the magnetic force? (c) If the direction of $\vec{v}\,'$ is reversed, so both charges are moving in the same direction, what are the magnitude and direction of the magnetic forces that the two charges exert on each other?

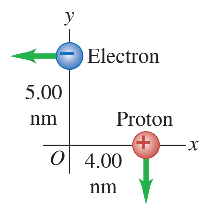
Figure **E28.6**



28.7 • At one instant, point P is 6.60 μm to the left of a proton that is moving at 3.30 km/s in vacuum. (a) What is the direction of the unit vector in Eq. (28.2)? (b) What is the magnetic field caused by this proton if it is moving to the right or to the left? (c) What is the magnetic field (magnitude and direction) caused by this proton if it is moving toward the top of the page? (d) What is the answer to part (c) for an electron instead of a proton?

28.8 •• An electron and a proton are each moving at 735 km/s in perpendicular paths as shown in **Fig. E28.8**. At the instant when they are at the positions shown, find the magnitude and direction of (a) the total magnetic field they produce at the origin; (b) the magnetic field the electron produces at the location of the proton; (c) the total electric force and the total magnetic force that the electron exerts on the proton.

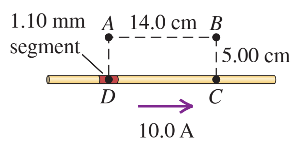
Figure **E28.8**



Section 28.2 Magnetic Field of a Current Element

28.9 • A straight wire carries a 10.0 A current (**Fig. E28.9**). *ABCD* is a rectangle with point *D* in the middle of a 1.10 mm segment of the wire and point *C* in the wire. Find the magnitude and direction of the magnetic field due to this segment at (a) point *A*; (b) point *B*; (c) point *C*.

Figure **E28.9**



28.10 • A short current element $d\vec{l} = (0.500 \text{ mm})\hat{\jmath}$ carries a current of 5.40 A in the same direction as $d\vec{l}$. Point *P* is located at $\vec{r} = (-0.730 \text{ m})\hat{\imath} + (0.390 \text{ m})\hat{k}$. Use unit vectors to express the magnetic field at *P* produced by this current element.

28.11 •• A long, straight wire lies along the *z*-axis and carries a 4.00 A current in the $+z$-direction. Find the magnetic field (magnitude and direction) produced at the following points by a 0.500 mm segment of the wire centered at the origin: (a) $x = 2.00$ m, $y = 0$, $z = 0$; (b) $x = 0$, $y = 2.00$ m, $z = 0$; (c) $x = 2.00$ m, $y = 2.00$ m, $z = 0$; (d) $x = 0$, $y = 0$, $z = 2.00$ m,

28.12 •• Two parallel wires are 5.00 cm apart and carry currents in opposite directions, as shown in **Fig. E28.12**. Find the magnitude and direction of the magnetic field at point *P* due to two 1.50 mm segments of wire that are opposite each other and each 8.00 cm from *P*.

Figure **E28.12**

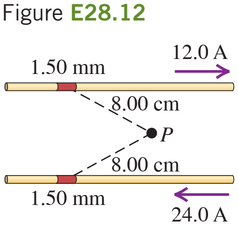
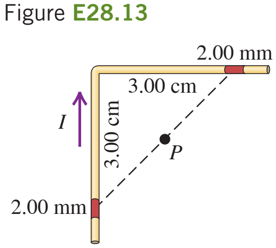


Figure **E28.13**



28.13 • A wire carrying a 28.0 A current bends through a right angle. Consider two 2.00 mm segments of wire, each 3.00 cm from the bend (**Fig. E28.13**). Find the magnitude and direction of the magnetic field these two segments produce at point *P*, which is midway between them.

Section 28.3 Magnetic Field of a Straight Current-Carrying Conductor

28.14 •• A long, straight wire lies along the *x*-axis and carries current $I = 60.0$ A in the $+x$-direction. A small particle with mass 3.00×10^{-6} kg and charge 8.00×10^{-3} C is traveling in the vicinity of the wire. At an instant when the particle is on the *y*-axis at $y = 8.00$ cm, its acceleration has components $a_x = -5.00 \times 10^3 \text{ m/s}^2$ and $a_y = +9.00 \times 10^3 \text{ m/s}^2$. At that instant what are the *x*- and *y*-components of the velocity of the particle?

28.15 • **The Magnetic Field from a Lightning Bolt.** Lightning bolts can carry currents up to approximately 20 kA. We can model such a current as the equivalent of a very long, straight wire. (a) If you were unfortunate enough to be 5.0 m away from such a lightning bolt, how large a magnetic field would you experience? (b) How does this field compare to one you would experience by being 5.0 cm from a long, straight household current of 10 A?

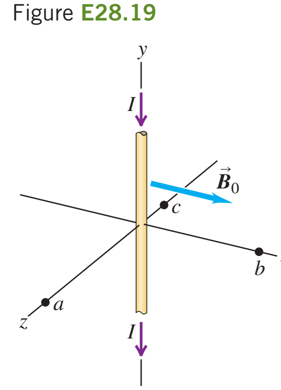
28.16 • A very long, straight horizontal wire carries a current such that 8.20×10^{18} electrons per second pass any given point going from west to east. What are the magnitude and direction of the magnetic field this wire produces at a point 4.00 cm directly above it?

28.17 • **BIO Currents in the Heart.** The body contains many small currents caused by the motion of ions in the organs and cells. Measurements of the magnetic field around the chest due to currents in the heart give values of about 10 μG. Although the actual currents are rather complicated, we can gain a rough understanding of their magnitude if we model them as a long, straight wire. If the surface of the chest is 5.0 cm from this current, how large is the current in the heart?

28.18 • Two long, straight wires, one above the other, are separated by a distance $2a$ and are parallel to the *x*-axis. Let the $+y$-axis be in the plane of the wires in the direction from the lower wire to the upper wire. Each wire carries current *I* in the $+x$-direction. What are the magnitude and direction of the net magnetic field of the two wires at a point in the plane of the wires (a) midway between them; (b) at a distance *a* above the upper wire; (c) at a distance *a* below the lower wire?

28.19 •• A long, straight wire lies along the *y*-axis and carries a current $I = 8.00$ A in the $-y$-direction (**Fig. E28.19**). In addition to the magnetic field due to the current in the wire, a uniform magnetic field $\vec{B}_0$ with magnitude 1.50×10^{-6} T is in the $+x$-direction. What is the total field (magnitude and direction) at the following points in the *xz*-plane: (a) $x = 0$, $z = 1.00$ m; (b) $x = 1.00$ m, $z = 0$; (c) $x = 0$, $z = -0.25$ m?

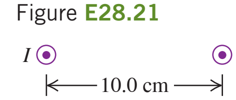
Figure **E28.19**



28.20 •• **BIO Transmission Lines and Health.** Currents in dc transmission lines can be 100 A or higher. Some people are concerned that the electromagnetic fields from such lines near their homes could pose health dangers. For a line that has current 150 A and a height of 8.0 m above the ground, what magnetic field does the line produce at ground level? Express your answer in teslas and as a percentage of the earth's magnetic field, which is 0.50 G. Is this value cause for worry?

28.21 • Two long, straight, parallel wires, 10.0 cm apart, carry equal 4.00 A currents in the same direction, as shown in **Fig. E28.21**. Find the magnitude and direction of the magnetic field at (a) point P_1, midway between the wires; (b) point P_2, 25.0 cm to the right of P_1; (c) point P_3, 20.0 cm directly above P_1.

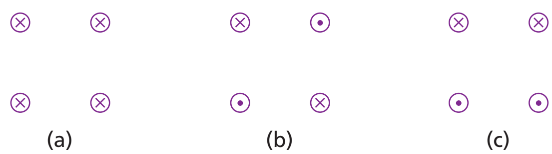
Figure **E28.21**



28.22 •• A rectangular loop with dimensions 4.20 cm by 9.50 cm carries current *I*. The current in the loop produces a magnetic field at the center of the loop that has magnitude 5.50×10^{-5} T and direction away from you as you view the plane of the loop. What are the magnitude and direction (clockwise or counterclockwise) of the current in the loop?

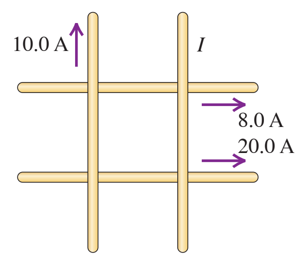
28.23 • Four long, parallel power lines each carry 100 A currents. A cross-sectional diagram of these lines is a square, 20.0 cm on each side. For each of the three cases shown in **Fig. E28.23**, calculate the magnetic field at the center of the square.

Figure **E28.23**



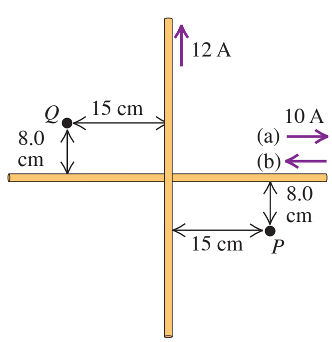
28.24 • Four very long, current-carrying wires in the same plane intersect to form a square 40.0 cm on each side, as shown in **Fig. E28.24**. Find the magnitude and direction of the current I so that the magnetic field at the center of the square is zero.

Figure **E28.24**



28.25 •• Two very long insulated wires perpendicular to each other in the same plane carry currents as shown in **Fig. E28.25**. Find the magnitude of the *net* magnetic field these wires produce at points P and Q if the 10.0 A current is (a) to the right or (b) to the left.

Figure **E28.25**

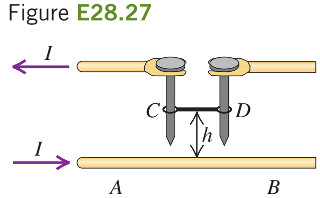


28.26 • A wire of length 20.0 cm lies along the x-axis with the center of the wire at the origin. The wire carries current $I = 8.00$ A in the $-x$-direction. (a) What is the magnitude B of the magnetic field of the wire at the point $y = 5.00$ cm on the y-axis? (b) What is the percent difference between the answer in (a) and the value you obtain if you assume the wire is infinitely long and use Eq. (28.9) to calculate B?

Section 28.4 Force Between Parallel Conductors

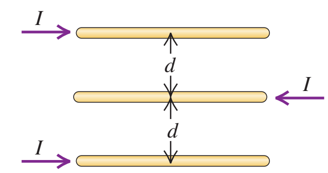
28.27 • A long, horizontal wire AB rests on the surface of a table and carries a current I. Horizontal wire CD is vertically above wire AB and is free to slide up and down on the two vertical metal guides C and D (**Fig. E28.27**). Wire CD is connected through the sliding contacts to another wire that also carries a current I, opposite in direction to the current in wire AB. The mass per unit length of the wire CD is λ. To what equilibrium height h will the wire CD rise, assuming that the magnetic force on it is due entirely to the current in the wire AB?

Figure **E28.27**



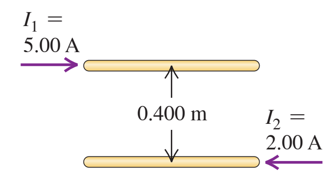
28.28 • Three very long parallel wires each carry current I in the directions shown in **Fig. E28.28**. If the separation between adjacent wires is d, calculate the magnitude and direction of the net magnetic force per unit length on each wire.

Figure **E28.28**



28.29 • Two long, parallel wires are separated by a distance of 0.400 m (**Fig. E28.29**). The currents I_1 and I_2 have the directions shown. (a) Calculate the magnitude of the force exerted by each wire on a 1.20 m length of the other. Is the force attractive or repulsive? (b) Each current is doubled, so that I_1 becomes 10.0 A and I_2 becomes 4.00 A. Now what is the magnitude of the force that each wire exerts on a 1.20 m length of the other?

Figure **E28.29**



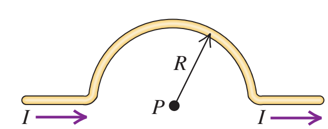
28.30 • Two long, parallel wires are separated by a distance of 2.50 cm. The force per unit length that each wire exerts on the other is 4.00×10^{-5} N/m, and the wires repel each other. The current in one wire is 0.600 A. (a) What is the current in the second wire? (b) Are the two currents in the same direction or in opposite directions?

Section 28.5 Magnetic Field of a Circular Current Loop

28.31 • **BIO Currents in the Brain.** The magnetic field around the head has been measured to be approximately 3.0×10^{-8} G. Although the currents that cause this field are quite complicated, we can get a rough estimate of their size by modeling them as a single circular current loop 16 cm (the width of a typical head) in diameter. What is the current needed to produce such a field at the center of the loop?

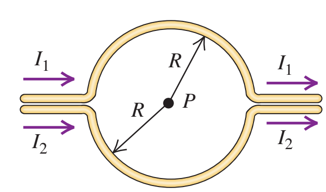
28.32 • Calculate the magnitude and direction of the magnetic field at point P due to the current in the semicircular section of wire shown in **Fig. E28.32**. (*Hint:* Does the current in the long, straight section of the wire produce any field at P?)

Figure **E28.32**



28.33 •• Calculate the magnitude of the magnetic field at point P of **Fig. E28.33** in terms of R, I_1, and I_2. What does your expression give when $I_1 = I_2$?

Figure **E28.33**



28.34 •• A closely wound, circular coil with radius 2.40 cm has 800 turns. (a) What must the current in the coil be if the magnetic field at the center of the coil is 0.0770 T? (b) At what distance x from the center of the coil, on the axis of the coil, is the magnetic field half its value at the center?

28.35 •• Two concentric circular loops of wire lie on a tabletop, one inside the other. The inner wire has a diameter of 20.0 cm and carries a clockwise current of 12.0 A, as viewed from above, and the outer wire has a diameter of 30.0 cm. What must be the magnitude and direction (as viewed from above) of the current in the outer wire so that the net magnetic field due to this combination of wires is zero at the common center of the wires?

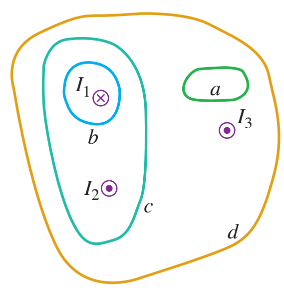
28.36 •• A closely wound coil has a radius of 6.00 cm and carries a current of 2.50 A. How many turns must it have if, at a point on the coil axis 6.00 cm from the center of the coil, the magnetic field is 6.39×10^{-4} T?

Section 28.6 Ampere's Law

28.37 • A closed curve encircles several conductors. The line integral $\oint \vec{B} \cdot d\vec{l}$ around this curve is $3.83 \times 10^{-4}\ \mathrm{T \cdot m}$. (a) What is the net current in the conductors? (b) If you were to integrate around the curve in the opposite direction, what would be the value of the line integral? Explain.

28.38 • **Figure E28.38** shows, in cross section, several conductors that carry currents through the plane of the figure. The currents have the magnitudes $I_1 = 4.0$ A, $I_2 = 6.0$ A, and $I_3 = 2.0$ A, and the directions shown. Four paths, labeled a through d, are shown. What is the line integral $\oint \vec{B} \cdot d\vec{l}$ for each path? Each integral involves going around the path in the counterclockwise direction. Explain your answers.

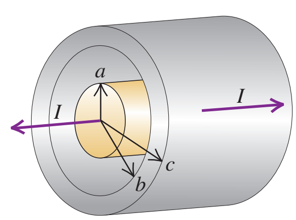
Figure E28.38



Section 28.7 Applications of Ampere's Law

28.39 • **Coaxial Cable.** A solid conductor with radius a is supported by insulating disks on the axis of a conducting tube with inner radius b and outer radius c (**Fig. E28.39**). The central conductor and tube carry equal currents I in opposite directions. The currents are distributed uniformly over the cross sections of each conductor. Derive an expression for the magnitude of the magnetic field (a) at points outside the central, solid conductor but inside the tube $(a < r < b)$ and (b) at points outside the tube $(r > c)$.

Figure E28.39



28.40 •• As a new electrical technician, you are designing a large solenoid to produce a uniform 0.150 T magnetic field near the center of the solenoid. You have enough wire for 4000 circular turns. This solenoid must be 55.0 cm long and 2.80 cm in diameter. What current will you need to produce the necessary field?

28.41 • Repeat Exercise 28.39 for the case in which the current in the central, solid conductor is I_1, the current in the tube is I_2, and these currents are in the same direction rather than in opposite directions.

28.42 •• A 15.0-cm-long solenoid with radius 0.750 cm is closely wound with 600 turns of wire. The current in the windings is 8.00 A. Compute the magnetic field at a point near the center of the solenoid.

28.43 •• A solenoid is designed to produce a magnetic field of 0.0270 T at its center. It has radius 1.40 cm and length 40.0 cm, and the wire can carry a maximum current of 12.0 A. (a) What minimum number of turns per unit length must the solenoid have? (b) What total length of wire is required?

28.44 • An ideal toroidal solenoid (see Example 28.10) has inner radius $r_1 = 15.0$ cm and outer radius $r_2 = 18.0$ cm. The solenoid has 250 turns and carries a current of 8.50 A. What is the magnitude of the magnetic field at the following distances from the center of the torus: (a) 12.0 cm; (b) 16.0 cm; (c) 20.0 cm?

28.45 • A magnetic field of 37.2 T has been achieved at the MIT Francis Bitter Magnet Laboratory. Find the current needed to achieve such a field (a) 2.00 cm from a long, straight wire; (b) at the center of a circular coil of radius 42.0 cm that has 100 turns; (c) near the center of a solenoid with radius 2.40 cm, length 32.0 cm, and 40,000 turns.

Section 28.8 Magnetic Materials

28.46 •• A toroidal solenoid with 400 turns of wire and a mean radius of 6.0 cm carries a current of 0.25 A. The relative permeability of the core is 80. (a) What is the magnetic field in the core? (b) What part of the magnetic field is due to the magnetic moments of the atoms in the core?

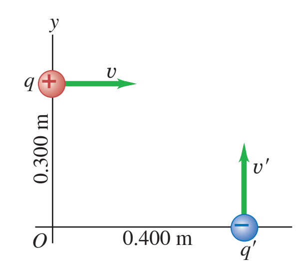
28.47 • A long solenoid with 60 turns of wire per centimeter carries a current of 0.15 A. The wire that makes up the solenoid is wrapped around a solid core of silicon steel ($K_m = 5200$). (The wire of the solenoid is jacketed with an insulator so that none of the current flows into the core.) (a) For a point inside the core, find the magnitudes of (i) the magnetic field $\vec{B}_0$ due to the solenoid current; (ii) the magnetization $\vec{M}$; (iii) the total magnetic field $\vec{B}$. (b) In a sketch of the solenoid and core, show the directions of the vectors $\vec{B}$, $\vec{B}_0$, and $\vec{M}$ inside the core.

28.48 • The current in the windings of a toroidal solenoid is 2.400 A. There are 500 turns, and the mean radius is 25.00 cm. The toroidal solenoid is filled with a magnetic material. The magnetic field inside the windings is found to be 1.940 T. Calculate (a) the relative permeability and (b) the magnetic susceptibility of the material that fills the toroid.

PROBLEMS

28.49 •• A pair of point charges, $q = +8.00\ \mu\mathrm{C}$ and $q' = -5.00\ \mu\mathrm{C}$, are moving as shown in **Fig. P28.49** with speeds $v = 9.00 \times 10^4$ m/s and $v' = 6.50 \times 10^4$ m/s. When the charges are at the locations shown in the figure, what are the magnitude and direction of (a) the magnetic field produced at the origin and (b) the magnetic force that q' exerts on q?

Figure P28.49



28.50 •• At a particular instant, charge $q_1 = +4.80 \times 10^{-6}$ C is at the point (0, 0.250 m, 0) and has velocity $\vec{v}_1 = (9.20 \times 10^5\ \mathrm{m/s})\hat{\imath}$. Charge $q_2 = -2.90 \times 10^{-6}$ C is at the point (0.150 m, 0, 0) and has velocity $\vec{v}_2 = (-5.30 \times 10^5\ \mathrm{m/s})\hat{\jmath}$. At this instant, what are the magnitude and direction of the magnetic force that q_1 exerts on q_2?

28.51 •• A long, straight wire lies along the x-axis and carries current $I_1 = 2.00$ A in the $+x$-direction. A second wire lies in the xy-plane and is parallel to the x-axis at $y = +0.800$ m. It carries current $I_2 = 6.00$ A, also in the $+x$-direction. In addition to $y \rightarrow \pm\infty$, at what point on the y-axis is the resultant magnetic field of the two wires equal to zero?

28.52 •• Repeat Problem 28.51 for I_2 in the $-x$-direction, with all the other quantities the same.

28.53 •• We can estimate the strength of the magnetic field of a refrigerator magnet in the following way: Imagine the magnet as a collection of current-loop magnetic dipoles. (a) Derive the force between two current loops with radius R and current I separated by distance $d \ll R$. Very close to the wire its magnetic field is about the same as for an infinitely long wire, and Eq.(28.11) can be used. (b) Using Eq. (28.17), express the current I in terms of the magnetic field at the middle of the loop, and express the radius R in terms of the area of the loop. In this way, derive an expression for the force F between two identical current loops separated by a small distance d in terms of their mutual area A and center magnetic field B. (c) Rearrange your result to obtain an expression for the magnetic field of a dipole with area A in terms of the force F from an identical dipole separated by a small distance d. (d) Now notice that the force it takes to separate one magnet from your refrigerator is nearly the same as the force it takes to separate two magnets stuck together. Estimate that force F. (e) Estimate the area of a refrigerator magnet. (f) Assume that when these magnets are stuck together or to the refrigerator, they are separated by an effective distance $d = 25\ \mu\mathrm{m}$. Use the formula derived above to estimate the magnetic field strength of the magnet.

28.54 • A long, straight wire carries a current of 8.60 A. An electron is traveling in the vicinity of the wire. At the instant when the electron is 4.50 cm from the wire and traveling at a speed of 6.00×10^4 m/s directly toward the wire, what are the magnitude and direction (relative to the direction of the current) of the force that the magnetic field of the current exerts on the electron?

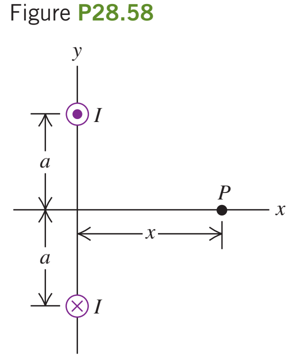
28.55 • **CP** A long, straight wire carries a 13.0 A current. An electron is fired parallel to this wire with a velocity of 250 km/s in the same direction as the current, 2.00 cm from the wire. (a) Find the magnitude and direction of the electron's initial acceleration. (b) What should be the magnitude and direction of a uniform electric field that will allow the electron to continue to travel parallel to the wire? (c) Is it necessary to include the effects of gravity? Justify your answer.

28.56 •• An electron is moving in the vicinity of a long, straight wire that lies along the x-axis. The wire has a constant current of 9.00 A in the $-x$-direction. At an instant when the electron is at point (0, 0.200 m, 0) and the electron's velocity is $\vec{v} = (5.00 \times 10^4 \text{ m/s})\hat{\imath} - (3.00 \times 10^4 \text{ m/s})\hat{\jmath}$, what is the force that the wire exerts on the electron? Express the force in terms of unit vectors, and calculate its magnitude.

28.57 •• **CP** (a) Determine the transmission power P of your cell phone. (This information is available online.) (b) A typical cell phone battery supplies a 1.5 V potential. If your phone battery supplies the power P, what is a good estimate of the current supplied by the battery? (c) Estimate the width of your head. (d) Estimate the diameter of the phone speaker that goes next to your ear. Model the current in the speaker as a current loop with the same diameter as the speaker. Use these values to estimate the magnetic field generated by your phone midway between the ears when it is held near one ear. (e) How does your answer compare to the earth's field, which is about 50 μT?

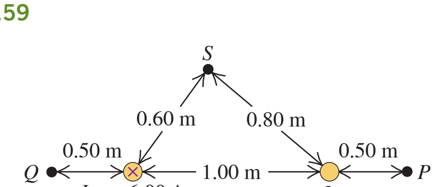
28.58 •• **Figure P28.58** shows an end view of two long, parallel wires perpendicular to the xy-plane, each carrying a current I but in opposite directions. (a) Copy the diagram, and draw vectors to show the $\vec{B}$ field of each wire and the net $\vec{B}$ field at point P. (b) Derive the expression for the magnitude of $\vec{B}$ at any point on the x-axis in terms of the x-coordinate of the point. What is the direction of $\vec{B}$? (c) Graph the magnitude of $\vec{B}$ at points on the x-axis. (d) At what value of x is the magnitude of $\vec{B}$ a maximum? (e) What is the magnitude of $\vec{B}$ when $x \gg a$?

Figure **P28.58**



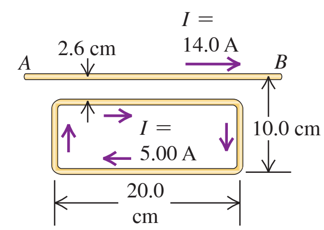
28.59 • Two long, straight, parallel wires are 1.00 m apart (**Fig. P28.59**). The wire on the left carries a current I_1 of 6.00 A into the plane of the paper. (a) What must the magnitude and direction of the current I_2 be for the net field at point P to be zero? (b) Then what are the magnitude and direction of the net field at Q? (c) Then what is the magnitude of the net field at S?

Figure **P28.59**



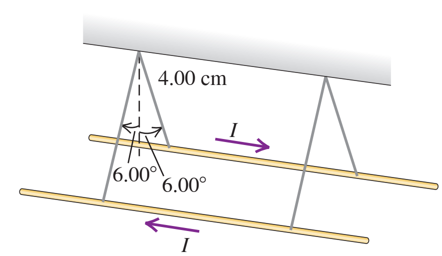
28.60 • The long, straight wire AB shown in **Fig. P28.60** carries a current of 14.0 A. The rectangular loop whose long edges are parallel to the wire carries a current of 5.00 A. Find the magnitude and direction of the net force exerted on the loop by the magnetic field of the wire.

Figure **P28.60**



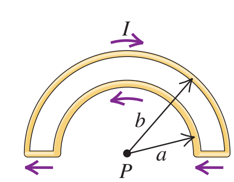
28.61 ••• **CP** Two long, parallel wires hang by 4.00-cm-long cords from a common axis (**Fig. P28.61**). The wires have a mass per unit length of 0.0125 kg/m and carry the same current in opposite directions. What is the current in each wire if the cords hang at an angle of 6.00° with the vertical?

Figure **P28.61**



28.62 • The wire semicircles shown in **Fig. P28.62** have radii a and b. Calculate the net magnetic field (magnitude and direction) that the current in the wires produces at point P.

Figure **P28.62**



28.63 • A long, straight, solid cylinder, oriented with its axis in the z-direction, carries a current whose current density is $\vec{J}$. The current density, although symmetric about the cylinder axis, is not constant and varies according to the relationship

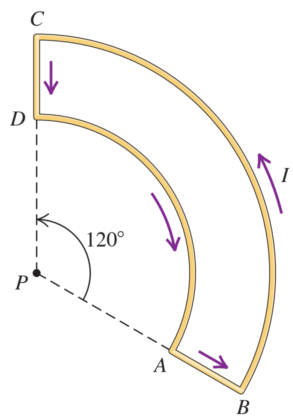
$$\vec{J} = \left(\frac{b}{r}\right) e^{(r-a)/\delta}\,\hat{k} \quad \text{for } r \le a$$

$$= 0 \quad \text{for } r \ge a$$

where the radius of the cylinder is $a = 5.00$ cm, r is the radial distance from the cylinder axis, b is a constant equal to 600 A/m, and δ is a constant equal to 2.50 cm. (a) Let I_0 be the total current passing through the entire cross section of the wire. Obtain an expression for I_0 in terms of b, δ, and a. Evaluate your expression to obtain a numerical value for I_0. (b) Using Ampere's law, derive an expression for the magnetic field $\vec{B}$ in the region $r \ge a$. Express your answer in terms of I_0 rather than b. (c) Obtain an expression for the current I contained in a circular cross section of radius $r \le a$ and centered at the cylinder axis. Express your answer in terms of I_0 rather than b. (d) Using Ampere's law, derive an expression for the magnetic field $\vec{B}$ in the region $r \le a$. (e) Evaluate the magnitude of the magnetic field at $r = \delta$, $r = a$, and $r = 2a$.

28.64 •• Calculate the magnetic field (magnitude and direction) at a point P due to a current $I = 12.0$ A in the wire shown in **Fig. P28.64**. Segment BC is an arc of a circle with radius 30.0 cm, and point P is at the center of curvature of the arc. Segment DA is an arc of a circle with radius 20.0 cm, and point P is at its center of curvature. Segments CD and AB are straight lines of length 10.0 cm each.

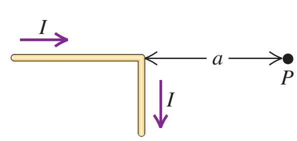
Figure **P28.64**



28.65 • CALC A long, straight wire with a circular cross section of radius R carries a current I. Assume that the current density is not constant across the cross section of the wire, but rather varies as $J = \alpha r$, where α is a constant. (a) By the requirement that J integrated over the cross section of the wire gives the total current I, calculate the constant α in terms of I and R. (b) Use Ampere's law to calculate the magnetic field $B(r)$ for (i) $r \leq R$ and (ii) $r \geq R$. Express your answers in terms of I.

28.66 • CALC The wire shown in **Fig. P28.66** is infinitely long and carries a current I. Calculate the magnitude and direction of the magnetic field that this current produces at point P.

Figure **P28.66**



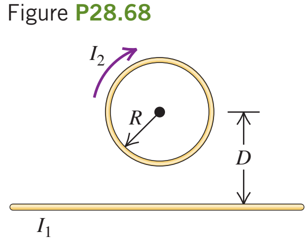
28.67 • CALC A long, straight, solid cylinder, oriented with its axis in the z-direction, carries a current whose current density is $\vec{J}$. The current density, although symmetric about the cylinder axis, is not constant but varies according to the relationship

$$\vec{J} = \frac{2I_0}{\pi a^2}\left[1 - \left(\frac{r}{a}\right)^2\right]\hat{k} \quad \text{for } r \leq a$$
$$= \mathbf{0} \quad \text{for } r \geq a$$

where a is the radius of the cylinder, r is the radial distance from the cylinder axis, and I_0 is a constant having units of amperes. (a) Show that I_0 is the total current passing through the entire cross section of the wire. (b) Using Ampere's law, derive an expression for the magnitude of the magnetic field $\vec{B}$ in the region $r \geq a$. (c) Obtain an expression for the current I contained in a circular cross section of radius $r \leq a$ and centered at the cylinder axis. (d) Using Ampere's law, derive an expression for the magnitude of the magnetic field $\vec{B}$ in the region $r \leq a$. How do your results in parts (b) and (d) compare for $r = a$?

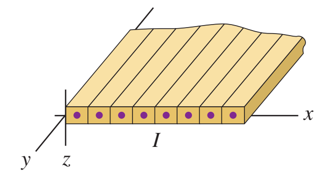
28.68 • A circular loop has radius R and carries current I_2 in a clockwise direction (**Fig. P28.68**). The center of the loop is a distance D above a long, straight wire. What are the magnitude and direction of the current I_1 in the wire if the magnetic field at the center of the loop is zero?

Figure **P28.68**



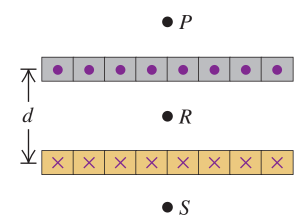
28.69 • An Infinite Current Sheet. Long, straight conductors with square cross sections and each carrying current I are laid side by side to form an infinite current sheet (**Fig. P28.69**). The conductors lie in the xy-plane, are parallel to the y-axis, and carry current in the $+y$-direction. There are n conductors per unit length measured along the x-axis. (a) What are the magnitude and direction of the magnetic field a distance a below the current sheet? (b) What are the magnitude and direction of the magnetic field a distance a above the current sheet?

Figure **P28.69**



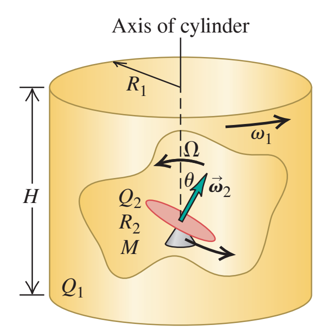
28.70 • Long, straight conductors with square cross section, each carrying current I, are laid side by side to form an infinite current sheet with current directed out of the plane of the page (**Fig. P28.70**). A second infinite current sheet is a distance d below the first and is parallel to it. The second sheet carries current into the plane of the page. Each sheet has n conductors per unit length. (Refer to Problem 28.69.) Calculate the magnitude and direction of the net magnetic field at (a) point P (above the upper sheet); (b) point R (midway between the two sheets); (c) point S (below the lower sheet).

Figure **P28.70**



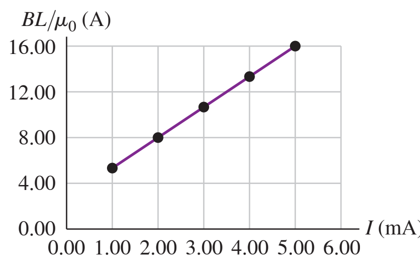
28.71 ••• CP A cylindrical shell with radius R_1 and height H has charge Q_1 and rotates around its axis with angular speed ω_1, as shown in **Fig. P28.71**. Inside the cylinder, far from its edges, sits a very small disk with radius R_2, mass M, and charge Q_2 mounted on a pivot, spinning with a large angular velocity $\vec{\omega}_2$ and oriented at angle θ with respect to the axis of the cylinder. The center of the disk is on the axis of the cylinder. The magnetic interaction between the cylinder and the disk causes a precession of the axis of the disk. (a) What is the magnitude of the enclosed current I_{encl} surrounded by a loop that has one vertical side that is along the axis of the cylinder and extends beyond the top and bottom of the cylinder? The other vertical side of the loop is very far outside the cylinder. (b) Assume the field is uniform within the cylinder and use Ampere's law to find the magnetic field at the center of the disk. (c) The magnetic moment of the disk has magnitude $\mu = \frac{1}{4}Q_2\omega_2 R_2^2$. What is the magnitude of the torque exerted on the disk? (d) What is the magnitude of the angular momentum of the disk?

Figure **P28.71**



28.72 •• DATA As a summer intern at a research lab, you are given a long solenoid that has two separate windings that are wound close together, in the same direction, on the same hollow cylindrical form. You must determine the number of turns in each winding. The solenoid has length $L = 40.0$ cm and diameter 2.80 cm. You let a 2.00 mA current flow in winding 1 and vary the current I in winding 2; both currents flow in the same direction. Then you measure the magnetic-field magnitude B at the center of the solenoid as a function of I. You plot your results as BL/μ_0 versus I. The graph in **Fig. P28.72** shows the best-fit straight line to your data. (a) Explain why the data plotted in this way should fall close to a straight line. (b) Use Fig. P28.72 to calculate N_1 and N_2, the number of turns in windings 1 and 2. (c) If the current in winding 1 remains 2.00 mA in its original direction and winding 2 has $I = 5.00$ mA in the opposite direction, what is B at the center of the solenoid?

Figure **P28.72**



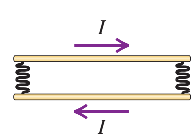
28.73 •• DATA You use a teslameter (a Hall-effect device) to measure the magnitude of the magnetic field at various distances from a long, straight, thick cylindrical copper cable that is carrying a large constant current. To exclude the earth's magnetic field from the measurement, you first set the meter to zero. You then measure the magnetic field B at distances x from the surface of the cable and obtain these data:

x (cm)	2.0	4.0	6.0	8.0	10.0
B (mT)	0.406	0.250	0.181	0.141	0.116

(a) You think you remember from your physics course that the magnetic field of a wire is inversely proportional to the distance from the wire. Therefore, you expect that the quantity Bx from your data will be constant. Calculate Bx for each data point in the table. Is Bx constant for this set of measurements? Explain. (b) Graph the data as x versus $1/B$. Explain why such a plot lies close to a straight line. (c) Use the graph in part (b) to calculate the current I in the cable and the radius R of the cable.

28.74 ••• DATA A pair of long, rigid metal rods, each of length 0.50 m, lie parallel to each other on a frictionless table. Their ends are connected by identical, very lightweight conducting springs with unstretched length l_0 and force constant k (**Fig. P28.74**). When a current I runs through the circuit consisting of the rods and springs, the springs stretch. You measure the distance x each spring stretches for certain values of I. When $I = 8.05$ A, you measure that $x = 0.40$ cm. When $I = 13.1$ A, you find $x = 0.80$ cm. In both cases the rods are much longer than the stretched springs, so it is accurate to use Eq. (28.11) for two infinitely long, parallel conductors. (a) From these two measurements, calculate l_0 and k. (b) If $I = 12.0$ A, what distance x will each spring stretch? (c) What current is required for each spring to stretch 1.00 cm?

Figure **P28.74**

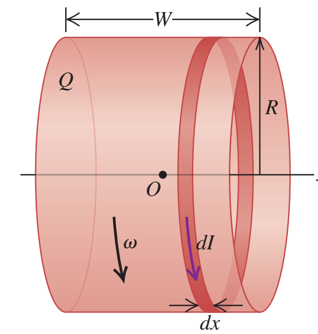


28.75 ••• CP CALC A plasma is a gas of ionized (charged) particles. When plasma is in motion, magnetic effects "squeeze" its volume, inducing inward pressure known as a pinch. Consider a cylindrical tube of plasma with radius R and length L moving with velocity $\vec{v}$ along its axis. If there are n ions per unit volume and each ion has charge q, we can determine the pressure felt by the walls of the cylinder. (a) What is the volume charge density ρ in terms of n and q? (b) The thickness of the cylinder "surface" is $n^{-1/3}$. What is the surface charge density σ in terms of n and q? (c) The current density inside the cylinder is $\vec{J} = \rho\vec{v}$. Use this result along with Ampere's law to determine the magnetic field on the surface of the cylinder. Denote the circumferential unit vector as $\hat{\phi}$. (d) The width of a differential strip of surface current is $R\,d\phi$. What is the differential current dI_{surface} that flows along this strip? (e) What differential force is felt by this strip due to the magnetic field generated by the volume current? (f) Integrate to determine the total force on the walls of the cylinder; then divide by the wall area to obtain the pressure in terms of n, q, R, and v. (g) If a plasma cylinder with radius 2.0 cm has a charge density of 8.0×10^{16} ions/cm^3, where each ion has a charge of $e = 1.6 \times 10^{-19}$ C and is moving axially with a speed of 20.0 m/s, what is its pinch pressure?

CHALLENGE PROBLEMS

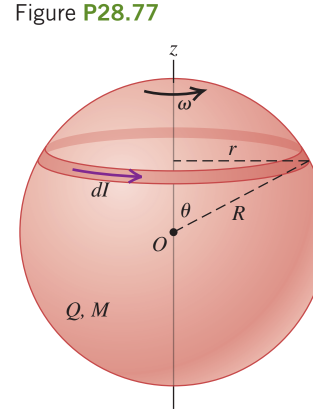
28.76 ••• CALC A cylindrical shell with radius R and length W carries a uniform charge Q and rotates about its axis with angular speed ω. The center of the cylinder lies at the origin O and its axis is coincident with the x-axis, as shown in **Fig. P28.76**. (a) What is the charge density σ? (b) What is the differential current dI on a circular strip of the cylinder centered at x and with width dx? (c) Use Eq. (28.15) to write an expression for the differential magnetic field $d\vec{B}$ at the origin due to this strip. (d) Integrate to determine the magnetic field at the origin.

Figure **P28.76**



28.77 ••• CALC When a rigid charge distribution with charge Q and mass M rotates about an axis, its magnetic moment $\vec{\mu}$ is linearly proportional to its angular momentum $\vec{L}$, with $\vec{\mu} = \gamma\vec{L}$. The constant of proportionality γ is called the gyromagnetic ratio of the object. We can write $\gamma = g(Q/2M)$, where g is a dimensionless number called the g-factor of the object. Consider a spherical shell with mass M and uniformly distributed charge Q centered on the origin O and rotating about the z-axis with angular speed ω. (a) A thin slice with latitude θ measured with respect to the positive z-axis describes a current loop with width $R\,d\theta$ and radius $r = R\sin\theta$, as shown in **Fig. P28.77**. What is the differential current dI carried by this loop, in terms of Q, ω, R, θ, and $d\theta$? (b) The differential magnetic moment contributed by that slice is $d\mu = A\,dI$, where $A = \pi r^2$ is the area enclosed by the loop. Express the differential magnetic moment in terms of Q, ω, R, θ, and $d\theta$. (c) Integrate over θ to determine the magnetic moment $\vec{\mu}$. (d) What is the magnitude of the angular momentum $\vec{L}$? (e) Determine the gyromagnetic ratio γ. (f) What is the g-factor for a spherical shell?

Figure **P28.77**



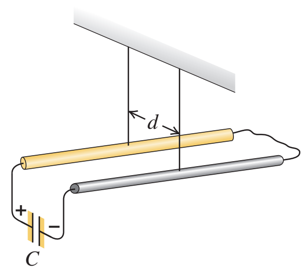
28.78 ••• CALC The law of Biot and Savart in Eq. (28.7) generalizes to the case of surface currents as

$$\vec{B} = \frac{\mu_0}{4\pi}\int \frac{\sigma\vec{v} \times \hat{r}}{r^2}\,da$$

where σ is the local charge density, $\vec{v}$ is the local velocity, and da is a differential area element. Re-visit Challenge Problem 28.76 and use the above equation as an alternative means to derive the magnetic field at the center of the cylinder. Use the following steps: (a) Write the charge density σ. (b) The origin is at the center of the cylinder. What is the vector $\vec{v}$ that points from the element with coordinates $(x, y, z) = (x, R\cos\phi, R\sin\phi)$ to the origin? (c) What is the velocity $\vec{v}$ of the element? (d) What is the vector product $\vec{v} \times \hat{r}$? (e) An area element on the cylinder may be written as $da = R\,dx\,d\phi$. Use this and the previously established information to write the generalized law of Biot and Savart as a double integral. Evaluate the integral to determine the magnetic field $\vec{B}$ at the center of the cylinder. (f) Is your result consistent with your result in Challenge Problem 28.76?

28.79 ••• **CP** Two long, straight conducting wires with linear mass density λ are suspended from cords so that they are each horizontal, parallel to each other, and a distance d apart. The back ends of the wires are connected to each other by a slack, low-resistance connecting wire. A charged capacitor (capacitance C) is now added to the system; the positive plate of the capacitor (initial charge $+Q_0$) is connected to the front end of one of the wires, and the negative plate of the capacitor (initial charge $-Q_0$) is connected to the front end of the other wire (**Fig. P28.79**). Both of these connections are also made by slack, low-resistance wires. When the connection is made, the wires are pushed aside by the repulsive force between the wires, and each wire has an initial horizontal velocity of magnitude v_0. Assume that the time constant for the capacitor to discharge is negligible compared to the time it takes for any appreciable displacement in the position of the wires to occur. (a) Show that the initial speed v_0 of either wire is given by

Figure **P28.79**



$$v_0 = \frac{\mu_0 Q_0^2}{4\pi\lambda RCd}$$

where R is the total resistance of the circuit. (b) To what height h will each wire rise as a result of the circuit connection?

28.80 ••• A wide, long, insulating belt has a uniform positive charge per unit area σ on its upper surface. Rollers at each end move the belt to the right at a constant speed v. Calculate the magnitude and direction of the magnetic field produced by the moving belt at a point just above its surface. (*Hint:* At points near the surface and far from its edges or ends, the moving belt can be considered to be an infinite current sheet like that in Problem 28.69.)

MCAT-STYLE PASSAGE PROBLEMS

BIO Studying Magnetic Bacteria. Some types of bacteria contain chains of ferromagnetic particles parallel to their long axis. The chains act like small bar magnets that align these *magnetotactic* bacteria with the earth's magnetic field. In one experiment to study the response of such bacteria to magnetic fields, a solenoid is constructed with copper wire 1.0 mm in diameter, evenly wound in a single layer to form a helical coil of length 40 cm and diameter 12 cm. The wire has a very thin layer of insulation, and the coil is wound so that adjacent turns are just touching. The solenoid, which generates a magnetic field, is in an enclosure that shields it from other magnetic fields. A sample of magnetotactic bacteria is placed inside the solenoid. The torque on an individual bacterium in the solenoid's magnetic field is proportional to the magnitude of the magnetic field and to the sine of the angle between the long axis of the bacterium and the magnetic-field direction.

28.81 What current is needed in the wire so that the magnetic field experienced by the bacteria has a magnitude of 150 μT? (a) 0.095 A; (b) 0.12 A; (c) 0.30 A; (d) 14 A.

28.82 To use a larger sample, the experimenters construct a solenoid that has the same length, type of wire, and loop spacing but twice the diameter of the original. How does the maximum possible magnetic torque on a bacterium in this new solenoid compare with the torque the bacterium would have experienced in the original solenoid? Assume that the currents in the solenoids are the same. The maximum torque in the new solenoid is (a) twice that in the original one; (b) half that in the original one; (c) the same as that in the original one; (d) one-quarter that in the original one.

28.83 The solenoid is removed from the enclosure and then used in a location where the earth's magnetic field is 50 μT and points horizontally. A sample of bacteria is placed in the center of the solenoid, and the same current is applied that produced a magnetic field of 150 μT in the lab. Describe the field experienced by the bacteria: The field (a) is still 150 μT; (b) is now 200 μT; (c) is between 100 and 200 μT, depending on how the solenoid is oriented; (d) is between 50 and 150 μT, depending on how the solenoid is oriented.

ANSWERS

Chapter Opening Question ?

(iv) There would be *no* change in the magnetic field strength. From Example 28.9 (Section 28.7), the field inside a solenoid has magnitude $B = \mu_0 nI$, where n is the number of turns of wire per unit length. Joining two solenoids end to end doubles both the number of turns and the length, so the number of turns per unit length is unchanged.

Key Example VARIATION Problems

VP28.2.1 (a) $(3.20 \times 10^{-15}\ \text{T})\hat{k}$ (b) $(-8.00 \times 10^{-16}\ \text{T})\hat{j}$ (c) zero (d) $(1.13 \times 10^{-15}\ \text{T})\hat{k}$

VP28.2.2 (a) 1.13×10^{-28} N in the $-x$-direction (toward the proton) (b) 1.13×10^{-28} N in the $+x$-direction (toward the electron)

VP28.2.3 (a) $(0.600)\hat{\imath} + (0.800)\hat{j}$ (b) $(-2.88 \times 10^{-11}\ \text{T})\hat{k}$

VP28.2.4 $(9.83 \times 10^{-23}\ \text{N})\hat{k}$

VP28.5.1 2.24 A

VP28.5.2 (a) $(-6.00 \times 10^{-5}\ \text{T})\hat{\imath}$ (b) $(-4.67 \times 10^{-5}\ \text{T})\hat{\imath}$ (c) $(8.67 \times 10^{-5}\ \text{T})\hat{\imath}$

VP28.5.3 (a) 5.00 A (b) 1.00 A

VP28.5.4 3.29×10^4 A, upward

VP28.10.1 (a) 4.94×10^{-6} T (b) 8.89×10^{-6} T (c) 6.67×10^{-6} T

VP28.10.2 (a) $B = 0$ for $r < R_1$, $B = (\mu_0 J/2r)(r^2 - R_1^2)$ for $R_2 > r > R_1$, $B = (\mu_0 J/2r)(R_2^2 - R_1^2)$ for $r > R_2$ (b) $r = R_2$

VP28.10.3 3.18 A

VP28.10.4 (a) 833 (b) 2.33×10^{-3} T maximum, 1.75×10^{-3} T minimum

Bridging Problem

$$B = \frac{\mu_0 nQ}{a}$$

? The card reader at your bank's cash machine scans the information that is coded in a magnetic pattern on the back of your card. Why must you remove the card quickly rather than hold it motionless in the card reader's slot? (i) To maximize the magnetic force on the card; (ii) to maximize the magnetic force on the mobile charges in the card reader; (iii) to generate an electric force on the card; (iv) to generate an electric force on the mobile charges in the card reader.

29 Electromagnetic Induction

Almost every modern device or machine, from a computer to a washing machine to a power drill, has electric circuits at its heart. We learned in Chapter 25 that an electromotive force (emf) is required for a current to flow in a circuit; in Chapters 25 and 26 we almost always took the source of emf to be a battery. But for most devices that you plug into a wall socket, the source of emf is *not* a battery but an electric generating station. Such a station produces electrical energy by converting other forms of energy: gravitational potential energy at a hydroelectric plant, chemical energy in a coal-, gas-, or oil-fired plant, nuclear energy at a nuclear plant. But how is this energy conversion done?

The answer is a phenomenon known as *electromagnetic induction:* If the magnetic flux through a circuit changes, an emf and a current are induced in the circuit. In a power-generating station, magnets move relative to coils of wire to produce a changing magnetic flux in the coils and hence an emf.

The central principle of electromagnetic induction is *Faraday's law.* This law relates induced emf to changing magnetic flux in any loop, including a closed circuit. We also discuss Lenz's law, which helps us to predict the directions of induced emfs and currents. These principles will allow us to understand electrical energy-conversion devices such as motors, generators, and transformers.

Electromagnetic induction tells us that a time-varying magnetic field can act as a source of electric field. We'll also see how a time-varying *electric* field can act as a source of *magnetic* field. These remarkable results form part of a neat package of formulas, called *Maxwell's equations,* that describe the behavior of electric and magnetic fields in general. Maxwell's equations pave the way toward an understanding of electromagnetic waves, the topic of Chapter 32.

LEARNING OUTCOMES

In this chapter, you'll learn...

29.1 The experimental evidence that a changing magnetic field induces an emf.
29.2 How Faraday's law relates the induced emf in a loop to the change in magnetic flux through the loop.
29.3 How to determine the direction of an induced emf.
29.4 How to calculate the emf induced in a conductor moving through a magnetic field.
29.5 How a changing magnetic flux generates a circulating electric field.
29.6 How eddy currents arise in a metal that moves in a magnetic field.
29.7 The four fundamental equations that completely describe both electricity and magnetism.
29.8 The remarkable electric and magnetic properties of superconductors.

You'll need to review...

23.1 Conservative electric fields.
25.4 Electromotive force (emf).
27.3, 27.8, 27.9 Magnetic flux; direct-current motors; Hall effect.
28.5–28.7 Magnetic field of a current loop and solenoid; Ampere's law.

29.1 INDUCTION EXPERIMENTS

During the 1830s, several pioneering experiments with magnetically induced emf were carried out in England by Michael Faraday and in the United States by Joseph Henry (1797–1878). **Figure 29.1** shows several examples. In Fig. 29.1a, a coil of wire is connected to a galvanometer. When the nearby magnet is stationary, the meter shows no current. This isn't surprising; there is no source of emf in the circuit. But when we *move* the magnet either toward or away from the coil, the meter shows current in the circuit, but *only* while the magnet is moving (Fig. 29.1b). If we keep the magnet stationary and move the coil, we again detect a current during the motion. We call this an **induced current,** and the corresponding emf required to cause this current is called an **induced emf.**

In Fig. 29.1c we replace the magnet with a second coil connected to a battery. When the second coil is stationary, there is no current in the first coil. However, when we move the second coil toward or away from the first or move the first toward or away from the second, there is current in the first coil, but again *only* while one coil is moving relative to the other.

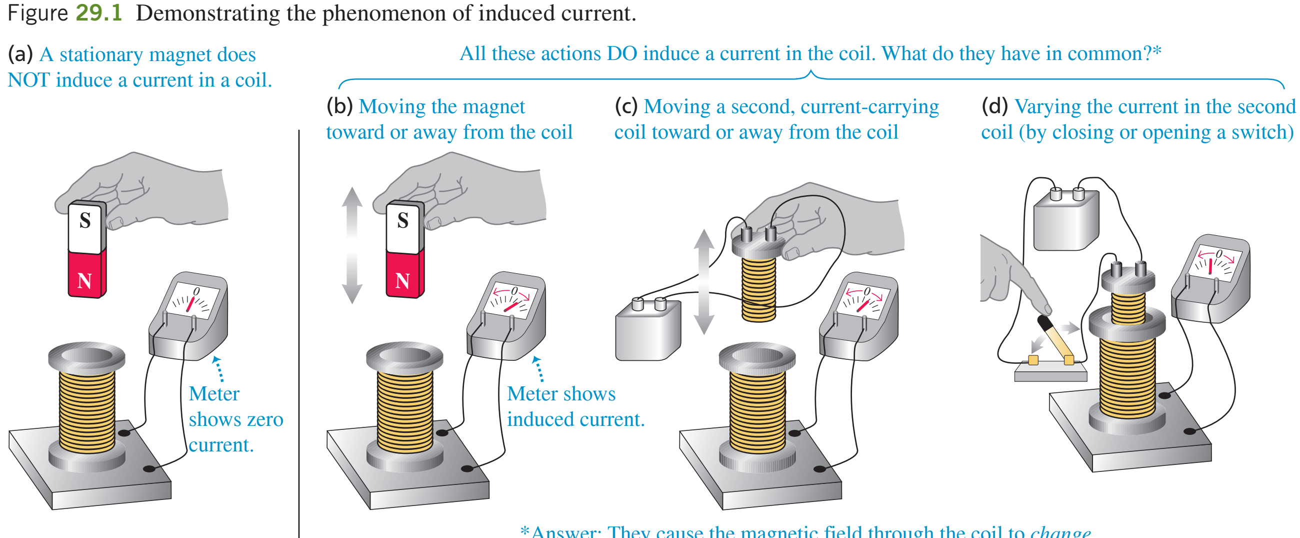
Finally, using the two-coil setup in Fig. 29.1d, we keep both coils stationary and vary the current in the second coil by opening and closing the switch. As we open or close the switch, there is a momentary current pulse in the first coil. The induced current in the first coil is present only while the current in the second coil is changing.

To explore further the common elements in these observations, let's consider a more detailed series of experiments (**Fig. 29.2**). We connect a coil of wire to a galvanometer and then place the coil between the poles of an electromagnet whose magnetic field we can vary. Here's what we observe:

1. When there is no current in the electromagnet, so that $\vec{B} = \mathbf{0}$, the galvanometer shows no current.
2. When the electromagnet is turned on, there is a momentary current through the meter as $\vec{B}$ increases.
3. When $\vec{B}$ levels off at a steady value, the current drops to zero.
4. With the coil in a horizontal plane, we squeeze it so as to decrease the cross-sectional area of the coil. The meter detects current only *during* the deformation, not before or after. When we increase the area to return the coil to its original shape, there is current in the opposite direction, but only while the area of the coil is changing.

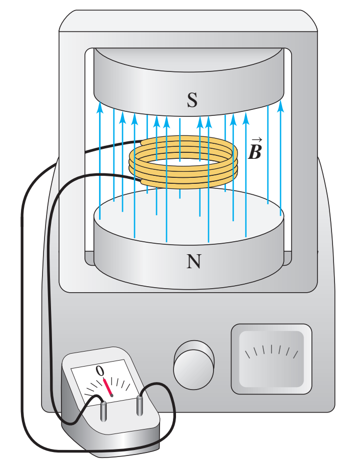
BIO APPLICATION Exploring the Brain with Induced emfs Transcranial magnetic stimulation (TMS) is a technique for studying the function of various parts of the brain. A coil held to the subject's head carries a varying electric current, and so produces a varying magnetic field. This field causes an induced emf, and that triggers electric activity in the region of the brain underneath the coil. By observing how the TMS subject responds (for instance, which muscles move as a result of stimulating a certain part of the brain), a physician can test for various neurological conditions.

Figure **29.1** Demonstrating the phenomenon of induced current.



5. If we rotate the coil a few degrees about a horizontal axis, the meter detects current during the rotation, in the same direction as when we decreased the area. When we rotate the coil back, there is a current in the opposite direction during this rotation.
6. If we jerk the coil out of the magnetic field, there is a current during the motion, in the same direction as when we decreased the area.
7. If we decrease the number of turns in the coil by unwinding one or more turns, there is a current during the unwinding, in the same direction as when we decreased the area. If we wind more turns onto the coil, there is a current in the opposite direction during the winding.
8. When the magnet is turned off, there is a momentary current in the direction opposite to the current when it was turned on.
9. The faster we carry out any of these changes, the greater the current.
10. If all these experiments are repeated with a coil that has the same shape but different material and different resistance, the current in each case is inversely proportional to the total circuit resistance. This shows that the induced emfs that are causing the current do not depend on the material of the coil but only on its shape and the magnetic field.

Figure **29.2** A coil in a magnetic field. When the $\vec{B}$ field is constant and the shape, location, and orientation of the coil do not change, no current is induced in the coil. A current is induced when any of these factors change.



The common element in these experiments is changing *magnetic flux* Φ_B through the coil connected to the galvanometer. In each case the flux changes either because the magnetic field changes with time or because the coil is moving through a nonuniform magnetic field. What's more, in each case the induced emf is proportional to the *rate of change* of magnetic flux Φ_B through the coil. The *direction* of the induced emf depends on whether the flux is increasing or decreasing. If the flux is constant, there is no induced emf.

Induced emfs have a tremendous number of practical applications. If you are reading these words indoors, you are making use of induced emfs right now! At the power plant that supplies your neighborhood, an electric generator produces an emf by varying the magnetic flux through coils of wire. (In the next section we'll see how this is done.) This emf supplies the voltage between the terminals of the wall sockets in your home, and this voltage supplies the power to your reading lamp.

Magnetically induced emfs, just like the emfs discussed in Section 25.4, are the result of *nonelectrostatic* forces. We have to distinguish carefully between the electrostatic electric fields produced by charges (according to Coulomb's law) and the nonelectrostatic electric fields produced by changing magnetic fields. We'll return to this distinction later in this chapter and the next.

29.2 FARADAY'S LAW

The common element in all induction effects is changing magnetic flux through a circuit. Before stating the simple physical law that summarizes all of the kinds of experiments described in Section 29.1, let's first review the concept of magnetic flux Φ_B (which we introduced in Section 27.3). For an infinitesimal-area element $d\vec{A}$ in a magnetic field $\vec{B}$ (**Fig. 29.3**), the magnetic flux $d\Phi_B$ through the area is

$$d\Phi_B = \vec{B} \cdot d\vec{A} = B_\perp \, dA = B \, dA \cos\phi$$

where $B_\perp$ is the component of $\vec{B}$ perpendicular to the surface of the area element and ϕ is the angle between $\vec{B}$ and $d\vec{A}$. (As in Chapter 27, be careful to distinguish between two quantities named "phi," ϕ and Φ_B.) The total magnetic flux Φ_B through a finite area is the integral of this expression over the area:

$$\Phi_B = \int \vec{B} \cdot d\vec{A} = \int B \, dA \cos\phi \qquad (29.1)$$

Figure **29.3** Calculating the magnetic flux through an area element.

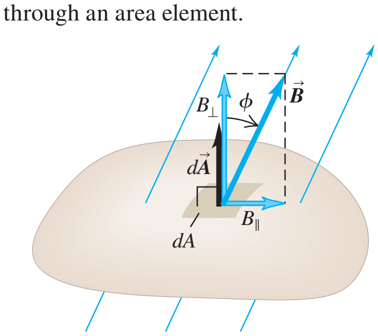
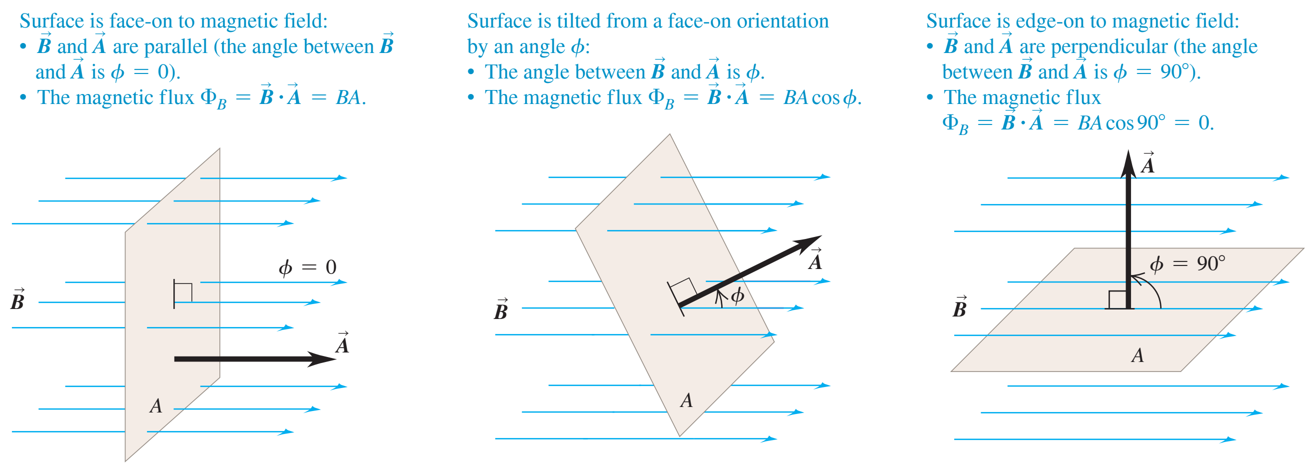


Figure 29.4 Calculating the flux of a uniform magnetic field through a flat area. (Compare to Fig. 22.6, which shows the rules for calculating the flux of a uniform *electric* field.)



If $\vec{B}$ is uniform over a flat area $\vec{A}$, then

$$\Phi_B = \vec{B} \cdot \vec{A} = BA\cos\phi \tag{29.2}$$

Figure 29.4 reviews the rules for using Eq. (29.2).

CAUTION Choosing the direction of $d\vec{A}$ or $\vec{A}$ In Eqs. (29.1) and (29.2) we must define the direction of the vector area $d\vec{A}$ or $\vec{A}$ unambiguously. There are always two directions perpendicular to any given area, and the sign of the magnetic flux through the area depends on which one we choose. For example, in Fig. 29.3 we chose $d\vec{A}$ to point upward, so ϕ is less than 90° and $\vec{B} \cdot d\vec{A}$ is positive. We could have chosen $d\vec{A}$ to point downward, in which case ϕ would have been greater than 90° and $\vec{B} \cdot d\vec{A}$ would have been negative. Both choices are equally good, but once we make a choice we must stick with it.

Faraday's law of induction states:

Faraday's law: The induced emf in a closed loop ... equals the negative of the time rate of change of magnetic flux through the loop.

$$\mathcal{E} = -\frac{d\Phi_B}{dt} \tag{29.3}$$

To understand the negative sign, we have to introduce a sign convention for the induced emf $\mathcal{E}$. But first let's look at a simple example of this law in action.

EXAMPLE 29.1 Emf and current induced in a loop

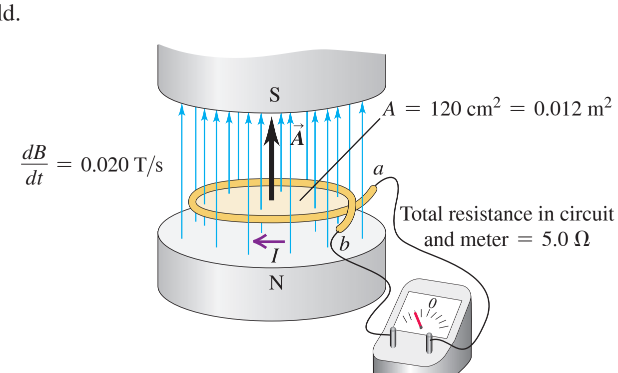
WITH VARIATION PROBLEMS

The magnetic field between the poles of the electromagnet in **Fig. 29.5** is uniform at any time, but its magnitude is increasing at the rate of 0.020 T/s. The area of the conducting loop in the field is 120 cm^2, and the total circuit resistance, including the meter, is 5.0 Ω. (a) Find the induced emf and the induced current in the circuit. (b) If the loop is replaced by one made of an insulator, what effect does this have on the induced emf and induced current?

IDENTIFY and SET UP The magnetic flux Φ_B through the loop changes as the magnetic field changes. Hence there will be an induced emf $\mathcal{E}$ and an induced current I in the loop. We calculate Φ_B from Eq. (29.2), then find $\mathcal{E}$ by using Faraday's law. Finally, we calculate I from $\mathcal{E} = IR$, where R is the total resistance of the circuit that includes the loop.

EXECUTE (a) The area vector $\vec{A}$ for the loop is perpendicular to the plane of the loop; we take $\vec{A}$ to be vertically upward. Then $\vec{A}$ and $\vec{B}$ are

Figure 29.5 A stationary conducting loop in an increasing magnetic field.



parallel, and because $\vec{B}$ is uniform the magnetic flux through the loop is $\Phi_B = \vec{B} \cdot \vec{A} = BA\cos 0 = BA$. The area $A = 0.012\ \text{m}^2$ is constant, so the rate of change of magnetic flux is

$$\frac{d\Phi_B}{dt} = \frac{d(BA)}{dt} = \frac{dB}{dt}A = (0.020\ \text{T/s})(0.012\ \text{m}^2)$$
$$= 2.4 \times 10^{-4}\ \text{V} = 0.24\ \text{mV}$$

This, apart from a sign that we haven't discussed yet, is the induced emf $\mathcal{E}$. The corresponding induced current is

$$I = \frac{\mathcal{E}}{R} = \frac{2.4 \times 10^{-4}\ \text{V}}{5.0\ \Omega} = 4.8 \times 10^{-5}\ \text{A} = 0.048\ \text{mA}$$

(b) By changing to an insulating loop, we've made the resistance of the loop very high. Faraday's law, Eq. (29.3), does not involve the resistance of the circuit in any way, so the induced *emf* does not change. But the *current* will be smaller, as given by the equation $I = \mathcal{E}/R$. If the loop is made of a perfect insulator with infinite resistance, the induced current is zero. This situation is analogous to an isolated battery whose terminals aren't connected to anything: An emf is present, but no current flows.

EVALUATE We can verify unit consistency in this calculation by noting that the magnetic-force relationship $\vec{F} = q\vec{v} \times \vec{B}$ implies that the units of $\vec{B}$ are the units of force divided by the units of (charge times velocity): $1\ \text{T} = (1\ \text{N})/(1\ \text{C} \cdot \text{m/s})$. The units of magnetic flux are then $(1\ \text{T})(1\ \text{m}^2) = 1\ \text{N} \cdot \text{s} \cdot \text{m/C}$, and the rate of change of magnetic flux is $1\ \text{N} \cdot \text{m/C} = 1\ \text{J/C} = 1\ \text{V}$. Thus the unit of $d\Phi_B/dt$ is the volt, as required by Eq. (29.3). Also recall from Section 27.3 that the unit of magnetic flux is the weber (Wb): $1\ \text{T} \cdot \text{m}^2 = 1\ \text{Wb}$, so $1\ \text{V} = 1\ \text{Wb/s}$.

KEYCONCEPT When there is a change in the magnetic flux through a curve in space that forms a closed loop, an emf is induced in that loop. If the loop is made of a conductor, an induced current flows in response to the emf.

Direction of Induced emf

We can find the direction of an induced emf or current by using Eq. (29.3) together with some simple sign rules. Here's the procedure:

1. Define a positive direction for the vector area $\vec{A}$.
2. From the directions of $\vec{A}$ and the magnetic field $\vec{B}$, determine the sign of the magnetic flux Φ_B and its rate of change $d\Phi_B/dt$. **Figure 29.6** shows several examples.
3. Determine the sign of the induced emf or current. If the flux is increasing, so $d\Phi_B/dt$ is positive, then the induced emf or current is negative; if the flux is decreasing, $d\Phi_B/dt$ is negative and the induced emf or current is positive.
4. Finally, use your right hand to determine the direction of the induced emf or current. Curl the fingers of your right hand around the $\vec{A}$ vector, with your right thumb in the direction of $\vec{A}$. If the induced emf or current in the circuit is *positive,* it is in the same direction as your curled fingers; if the induced emf or current is *negative,* it is in the opposite direction.

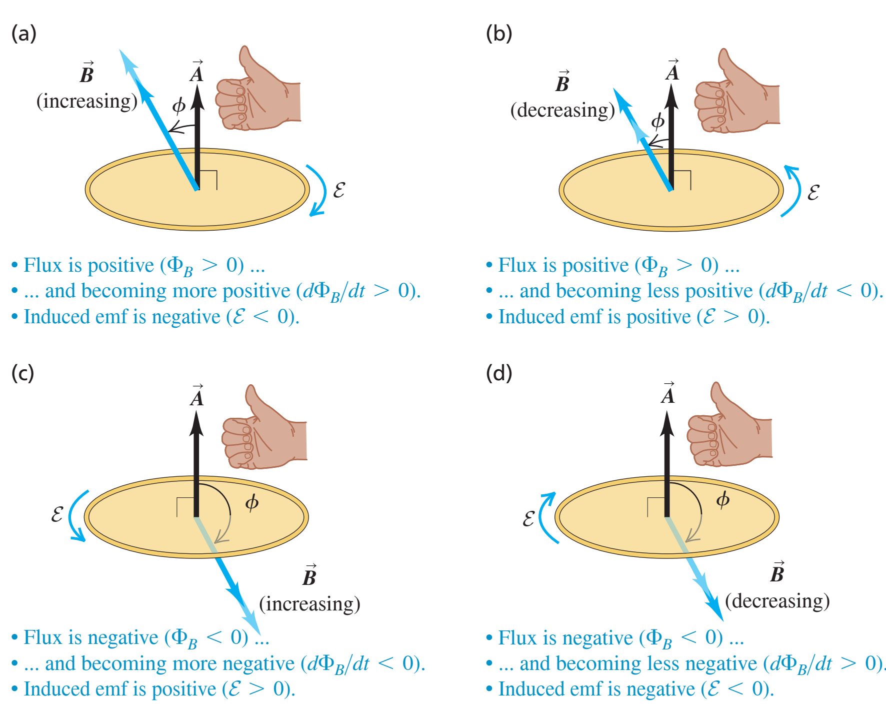


Figure **29.6** The magnetic flux is becoming (a) more positive, (b) less positive, (c) more negative, and (d) less negative. Therefore Φ_B is increasing in (a) and (d) and decreasing in (b) and (c). In (a) and (d) the emfs are negative (they are opposite to the direction of the curled fingers of your right hand when your right thumb points along $\vec{A}$). In (b) and (c) the emfs are positive (in the same direction as the curled fingers).

CAUTION Induced emfs are caused by changes in flux It can be tempting to think that *flux* is the cause of induced emf and that an induced emf will appear in a circuit whenever there is a magnetic field in the region bordered by the circuit. But Eq. (29.3) shows that only a *change* in flux through a circuit, not flux itself, can induce an emf in a circuit. If the flux through a circuit has a constant value, whether positive, negative, or zero, there is no induced emf. Note also that an induced emf will appear in a circuit if the flux is changed for any reason, including rotating the circuit relative to the magnetic field or changing the circuit's shape.

In Example 29.1, in which $\vec{A}$ is upward, a positive $\mathcal{E}$ would be directed counterclockwise around the loop, as seen from above. Both $\vec{A}$ and $\vec{B}$ are upward in this example, so Φ_B is positive; the magnitude B is increasing, so $d\Phi_B/dt$ is positive. Hence by Eq. (29.3), $\mathcal{E}$ in Example 29.1 is *negative*—that is, *clockwise* around the loop, as seen from above.

If the loop in Fig. 29.5 is a conductor, the clockwise induced emf causes a clockwise induced current. This induced current produces an additional magnetic field through the loop, and the right-hand rule described in Section 28.5 shows that this field is *opposite* in direction to the increasing field produced by the electromagnet. This is an example of a general rule called *Lenz's law,* which says that any induction effect tends to oppose the change that caused it; in this case the change is the increase in the flux of the electromagnet's field through the loop. (We'll study Lenz's law in detail in the next section.)

You should check out the signs of the induced emfs and currents for the list of experiments in Section 29.1. For example, when the loop in Fig. 29.2 is in a constant field and we tilt it or squeeze it to *decrease* the flux through it, the induced emf and current are counterclockwise, as seen from above.

If a coil has N identical turns and if the flux varies at the same rate through each turn, the *total* rate of change through all turns is N times that for a single turn. If Φ_B is the flux through each turn, the total emf in a coil with N turns is

$$\mathcal{E} = -N\frac{d\Phi_B}{dt} \tag{29.4}$$

As we discussed in this chapter's introduction, induced emfs play an essential role in the generation of electrical power for commercial use. Several of the following examples explore different methods of producing emfs by changing the flux through a circuit.

PROBLEM-SOLVING STRATEGY 29.1 Faraday's Law

IDENTIFY *the relevant concepts:* Faraday's law applies when a magnetic flux is changing. To use the law, identify an area through which there is a flux of magnetic field. This will usually be the area enclosed by a loop made of a conducting material (though not always—see part (b) of Example 29.1). Identify the target variables.

SET UP *the problem* using the following steps:

1. Faraday's law relates the induced emf to the rate of change of magnetic flux. To calculate this rate of change, you first have to understand what is making the flux change. Is the conductor moving or changing orientation? Is the magnetic field changing?
2. The area vector $\vec{A}$ (or $d\vec{A}$) must be perpendicular to the plane of the area. You always have two choices of its direction; for example, if the area is in a horizontal plane, $\vec{A}$ could point up or down. Choose a direction and use it throughout the problem.

EXECUTE *the solution* as follows:

1. Calculate the magnetic flux from Eq. (29.2) if $\vec{B}$ is uniform over the area of the loop or Eq. (29.1) if it isn't uniform. Remember the direction you chose for the area vector.
2. Calculate the induced emf from Eq. (29.3) or (if your conductor has N turns in a coil) Eq. (29.4). Apply the sign rule (described just after Example 29.1) to determine the positive direction of emf.
3. If the circuit resistance is known, you can calculate the magnitude of the induced current I by using $\mathcal{E} = IR$.

EVALUATE *your answer:* Check your results for the proper units, and double-check that you have properly implemented the sign rules for magnetic flux and induced emf.

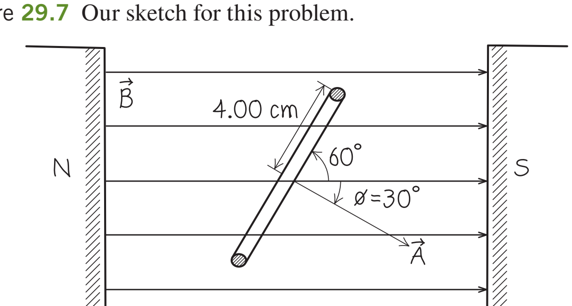
EXAMPLE 29.2 Magnitude and direction of an induced emf

WITH VARIATION PROBLEMS

A 500-loop circular wire coil with radius 4.00 cm is placed between the poles of a large electromagnet. The magnetic field is uniform and makes an angle of 60° with the plane of the coil; it decreases at 0.200 T/s. What are the magnitude and direction of the induced emf?

IDENTIFY and SET UP Our target variable is the emf induced by a varying magnetic flux through the coil. The flux varies because the magnetic field decreases in amplitude. We choose the area vector $\vec{A}$ to be in the direction shown in **Fig. 29.7**. With this choice, the geometry is similar to that of Fig. 29.6b. That figure will help us determine the direction of the induced emf.

Figure 29.7 Our sketch for this problem.



EXECUTE The magnetic field is uniform over the loop, so we can calculate the flux from Eq. (29.2): $\Phi_B = BA\cos\phi$, where $\phi = 30°$. In this expression, the only quantity that changes with time is the magnitude B of the field, so $d\Phi_B/dt = (dB/dt)A\cos\phi$.

CAUTION Remember how ϕ is defined You may have been tempted to say that $\phi = 60°$ in this problem. If so, remember that ϕ is the angle between $\vec{A}$ and $\vec{B}$, *not* the angle between $\vec{B}$ and the plane of the loop.

From Eq. (29.4), the induced emf in the coil of $N = 500$ turns is

$$\begin{aligned}\mathcal{E} &= -N\frac{d\Phi_B}{dt} = -N\frac{dB}{dt}A\cos\phi \\ &= -500(-0.200\ \text{T/s})\pi(0.0400\ \text{m})^2(\cos 30°) = 0.435\ \text{V}\end{aligned}$$

The positive answer means that when you point your right thumb in the direction of area vector $\vec{A}$ (30° below field $\vec{B}$ in Fig. 29.7), the positive direction for $\mathcal{E}$ is in the direction of the curled fingers of your right hand. If you viewed the coil from the left in Fig. 29.7 and looked in the direction of $\vec{A}$, the emf would be clockwise.

EVALUATE If the ends of the wire are connected, the direction of current in the coil is in the same direction as the emf—that is, clockwise as seen from the left side of the coil. A clockwise current increases the magnetic flux through the coil, and therefore tends to oppose the decrease in total flux. This is an example of Lenz's law, which we'll discuss in Section 29.3.

KEYCONCEPT Faraday's law states that the induced emf in a coil is proportional to the rate of change of the magnetic flux through the coil and to the number of turns in the coil. The sign of the induced emf tells you its direction.

EXAMPLE 29.3 Generator I: A simple alternator

WITH VARIATION PROBLEMS

Figure 29.8a shows a simple *alternator,* a device that generates an emf. A rectangular loop is rotated with constant angular speed ω about the axis shown. The magnetic field $\vec{B}$ is uniform and constant. At time $t = 0$, $\phi = 0$. Determine the induced emf.

IDENTIFY and SET UP The magnetic field $\vec{B}$ and the loop area A are constant, but the flux through the loop varies because the loop rotates and so the angle ϕ between $\vec{B}$ and the area vector $\vec{A}$ changes (Fig. 29.8a). Because the angular speed is constant and $\phi = 0$ at $t = 0$, the angle as a function of time is $\phi = \omega t$.

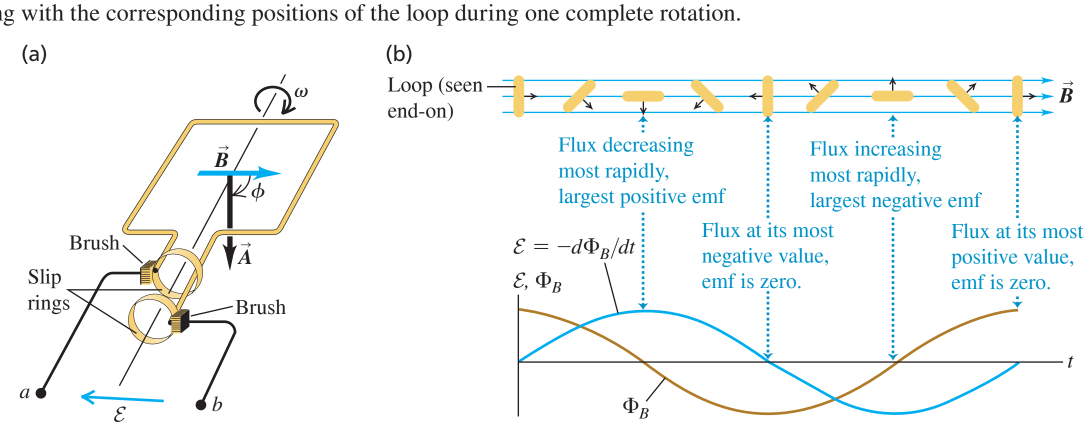
EXECUTE The magnetic field is uniform over the loop, so the magnetic flux is $\Phi_B = BA\cos\phi = BA\cos\omega t$. Hence, by Faraday's law [Eq. (29.3)] the induced emf is

$$\mathcal{E} = -\frac{d\Phi_B}{dt} = -\frac{d}{dt}(BA\cos\omega t) = \omega BA\sin\omega t$$

EVALUATE The induced emf $\mathcal{E}$ varies sinusoidally with time (see Fig. 29.8b). When the plane of the loop is perpendicular to $\vec{B}$ ($\phi = 0$ or 180°), Φ_B reaches its maximum and minimum values. At these times, its instantaneous rate of change is zero and $\mathcal{E}$ is zero. Conversely, $\mathcal{E}$ reaches its maximum and minimum values when the plane of the loop is parallel to $\vec{B}$ ($\phi = 90°$ or 270°) and Φ_B is changing most rapidly. We note that the induced emf does not depend on the *shape* of the loop, but only on its area.

We can use the alternator as a source of emf in an external circuit by using two *slip rings* that rotate with the loop, as shown in Fig. 29.8a. The rings slide against stationary contacts called *brushes,* which are connected to the output terminals a and b. Since the emf varies sinusoidally, the current that results in the circuit is an *alternating* current that also

Figure 29.8 (a) Schematic diagram of an alternator. A conducting loop rotates in a magnetic field, producing an emf. Connections from each end of the loop to the external circuit are made by means of that end's slip ring. The system is shown at the time when the angle $\phi = \omega t = 90°$. (b) Graph of the flux through the loop and the resulting emf between terminals a and b, along with the corresponding positions of the loop during one complete rotation.



Continued

varies sinusoidally in magnitude and direction. The amplitude of the emf can be increased by increasing the rotation speed, the field magnitude, or the loop area or by using N loops instead of one, as in Eq. (29.4).

Alternators are used in automobiles to generate the currents in the ignition, the lights, and the entertainment system. The arrangement is a little different than in this example; rather than having a rotating loop in a magnetic field, the loop stays fixed and an electromagnet rotates. (The rotation is provided by a mechanical connection between the alternator and the engine.) But the result is the same; the flux through the loop varies sinusoidally, producing a sinusoidally varying emf. Larger alternators of this same type are used in electric power plants (**Fig. 29.9**).

KEYCONCEPT In an alternator, a coil rotates relative to a magnetic field. The flux through the coil varies sinusoidally with time, as does the emf induced in the coil. The more rapid the rotation, the greater the magnitudes of the induced emf and current.

Figure **29.9** A commercial alternator uses many loops of wire wound around a barrel-like structure called an armature. The armature and wire remain stationary while electromagnets rotate on a shaft (not shown) through the center of the armature. The resulting induced emf is far larger than would be possible with a single loop of wire.

EXAMPLE 29.4 Generator II: A dc generator and back emf in a motor

The alternator in Example 29.3 produces a sinusoidally varying emf and hence an alternating current. **Figure 29.10a** shows a *direct-current* (dc) *generator* that produces an emf that always has the same sign. The arrangement of split rings, called a *commutator,* reverses the connections to the external circuit at angular positions at which the emf reverses. Figure 29.10b shows the resulting emf. Commercial dc generators have a large number of coils and commutator segments, smoothing out the bumps in the emf so that the terminal voltage is not only one-directional but also practically constant. This brush-and-commutator arrangement is the same as that in the direct-current motor discussed in Section 27.8. The motor's *back emf* is just the emf induced by the changing magnetic flux through its rotating coil. Consider a motor with a square, 500-turn coil 10.0 cm on a side. If the magnetic field has magnitude 0.200 T, at what rotation speed is the *average* back emf of the motor equal to 112 V?

IDENTIFY and SET UP As far as the rotating loop is concerned, the situation is the same as in Example 29.3 except that we now have N turns of wire. Without the commutator, the emf would alternate between positive and negative values and have an average value of zero (see Fig. 29.8b). With the commutator, the emf is never negative and its average value is positive (Fig. 29.10b). We'll use our result from Example 29.3 to obtain an expression for this average value and solve it for the rotational speed ω.

EXECUTE Comparison of Figs. 29.8b and 29.10b shows that the back emf of the motor is just N times the absolute value of the emf found for an alternator in Example 29.3, as in Eq. (29.4): $|\mathcal{E}| = N\omega BA|\sin\omega t|$. To find the *average* back emf, we must replace $|\sin\omega t|$ by its average value. We find this by integrating $|\sin\omega t|$ over half a cycle, from $t = 0$ to $t = T/2 = \pi/\omega$, and dividing by the elapsed time π/ω. During this half cycle, the sine function is positive, so $|\sin\omega t| = \sin\omega t$, and we find

$$(|\sin\omega t|)_{\text{av}} = \frac{\int_0^{\pi/\omega}\sin\omega t\,dt}{\pi/\omega} = \frac{2}{\pi}$$

The average back emf is then

$$\mathcal{E}_{\text{av}} = \frac{2N\omega BA}{\pi}$$

Solving for ω, we obtain

$$\omega = \frac{\pi\mathcal{E}_{\text{av}}}{2NBA} = \frac{\pi(112\text{ V})}{2(500)(0.200\text{ T})(0.100\text{ m})^2} = 176\text{ rad/s}$$

(Recall from Example 29.1 that $1\text{ V} = 1\text{ Wb/s} = 1\text{ T}\cdot\text{m}^2\text{/s}$.)

EVALUATE The average back emf is directly proportional to ω. Hence the slower the rotation speed, the less the back emf and the greater the possibility of burning out the motor, as we described in Example 27.11 (Section 27.8).

KEYCONCEPT The back emf in a motor is a result of Faraday's law. The average value of the back emf is proportional to the rotation speed of the motor.

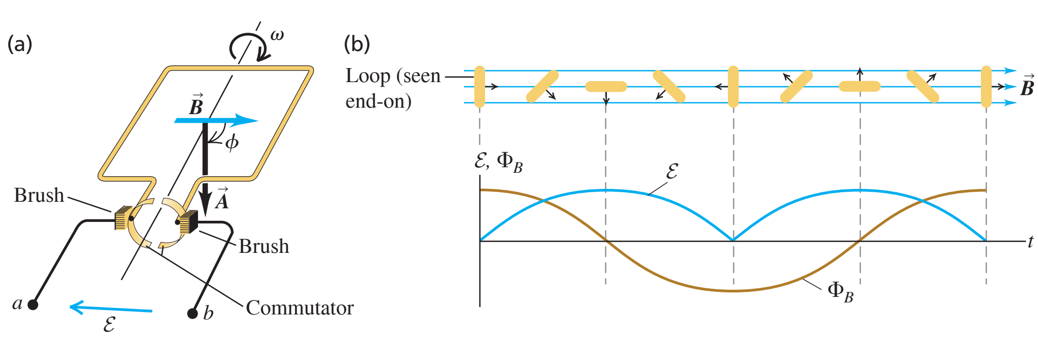


Figure **29.10** (a) Schematic diagram of a dc generator, using a split-ring commutator. The ring halves are attached to the loop and rotate with it. (b) Graph of the resulting induced emf between terminals a and b. Compare to Fig. 29.8b.

EXAMPLE 29.5 Generator III: The slidewire generator

Figure 29.11 shows a U-shaped conductor in a uniform magnetic field $\vec{B}$ perpendicular to the plane of the figure and directed *into* the page. We lay a metal rod (the "slidewire") with length L across the two arms of the conductor, forming a circuit, and move it to the right with constant velocity $\vec{v}$. This induces an emf and a current, which is why this device is called a *slidewire generator*. Find the magnitude and direction of the resulting induced emf.

IDENTIFY and SET UP The magnetic flux changes because the area of the loop—bounded on the right by the moving rod—is increasing. Our target variable is the emf $\mathcal{E}$ induced in this expanding loop. The magnetic field is uniform over the area of the loop, so we can find the flux from $\Phi_B = BA\cos\phi$. We choose the area vector $\vec{A}$ to point straight into the page, in the same direction as $\vec{B}$. With this choice a positive emf will be one that is directed clockwise around the loop. (You can check this with the right-hand rule: Using your right hand, point your thumb into the page and curl your fingers as in Fig. 29.6.)

EXECUTE Since $\vec{B}$ and $\vec{A}$ point in the same direction, the angle $\phi = 0$ and $\Phi_B = BA$. The magnetic-field magnitude B is constant, so the induced emf is

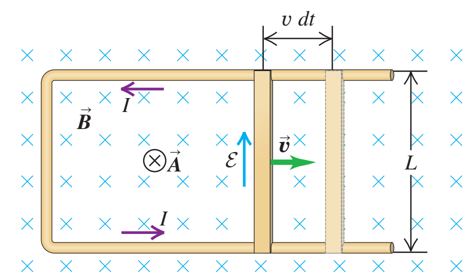
$$\mathcal{E} = -\frac{d\Phi_B}{dt} = -B\frac{dA}{dt}$$

To calculate dA/dt, note that in a time dt the sliding rod moves a distance $v\,dt$ (Fig. 29.11) and the loop area increases by an amount $dA = Lv\,dt$. Hence the induced emf is

$$\mathcal{E} = -B\frac{Lv\,dt}{dt} = -BLv$$

The minus sign tells us that the emf is directed *counterclockwise* around the loop. The induced current is also counterclockwise, as shown in the figure.

Figure **29.11** A slidewire generator. The magnetic field $\vec{B}$ and the vector area $\vec{A}$ are both directed into the figure. The increase in magnetic flux (caused by an increase in area) induces the emf and current.



EVALUATE The emf of a slidewire generator is constant if $\vec{v}$ is constant. Hence the slidewire generator is a *direct-current* generator. It's not a very practical device because the rod eventually moves beyond the U-shaped conductor and loses contact, after which the current stops.

KEYCONCEPT An induced emf can be produced if the area of a circuit in a magnetic field changes. The magnitude of the emf depends on the speed at which the area of the circuit changes.

EXAMPLE 29.6 Work and power in the slidewire generator

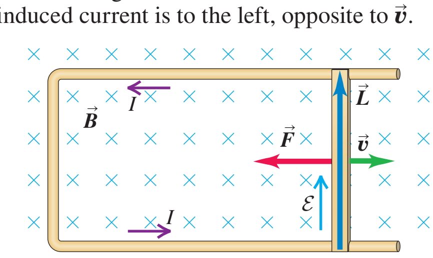
In the slidewire generator of Example 29.5, energy is dissipated in the circuit owing to its resistance. Let the resistance of the circuit (made up of the moving slidewire and the U-shaped conductor that connects the ends of the slidewire) at a given point in the slidewire's motion be R. Find the rate at which energy is dissipated in the circuit and the rate at which work must be done to move the rod through the magnetic field.

IDENTIFY and SET UP Our target variables are the *rates* at which energy is dissipated and at which work is done. Energy is dissipated in the circuit at the rate $P_{\text{dissipated}} = I^2R$. The current I in the circuit equals $|\mathcal{E}|/R$; we found an expression for the induced emf $\mathcal{E}$ in this circuit in Example 29.5. There is a magnetic force $\vec{F} = I\vec{L}\times\vec{B}$ on the rod, where $\vec{L}$ points along the rod in the direction of the current. **Figure 29.12** shows that this force is opposite to the rod velocity $\vec{v}$; to maintain the motion, whoever is pushing the rod must apply a force of equal magnitude in the direction of $\vec{v}$. This force does work at the rate $P_{\text{applied}} = Fv$.

EXECUTE First we'll calculate $P_{\text{dissipated}}$. From Example 29.5, $\mathcal{E} = -BLv$, so the current in the rod is $I = |\mathcal{E}|/R = Blv/R$. Hence

$$P_{\text{dissipated}} = I^2R = \left(\frac{BLv}{R}\right)^2 R = \frac{B^2L^2v^2}{R}$$

Figure **29.12** The magnetic force $\vec{F} = I\vec{L}\times\vec{B}$ that acts on the rod due to the induced current is to the left, opposite to $\vec{v}$.



To calculate P_{applied}, we first calculate the magnitude of $\vec{F} = I\vec{L}\times\vec{B}$. Since $\vec{L}$ and $\vec{B}$ are perpendicular, this magnitude is

$$F = ILB = \frac{BLv}{R}LB = \frac{B^2L^2v}{R}$$

The applied force has the same magnitude and does work at the rate

$$P_{\text{applied}} = Fv = \frac{B^2L^2v^2}{R}$$

Continued

CAUTION **You can't violate energy conservation** You might think that reversing the direction of $\vec{B}$ or of $\vec{v}$ would allow the magnetic force $\vec{F} = I\vec{L} \times \vec{B}$ to be in the *same* direction as $\vec{v}$. This would be a neat trick. Once the rod was moving, the changing magnetic flux would induce an emf and a current, and the magnetic force on the rod would make it move even faster, increasing the emf and current until the rod was moving at tremendous speed and producing electrical power at a prodigious rate. If this seems too good to be true and a violation of energy conservation, that's because it is. Reversing $\vec{B}$ also reverses the sign of the induced emf and current and hence the direction of $\vec{L}$, so the magnetic force still opposes the motion of the rod; a similar result holds true if we reverse $\vec{v}$.

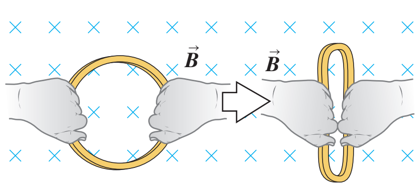
EVALUATE The rate at which work is done is exactly *equal* to the rate at which energy is dissipated in the resistance.

KEYCONCEPT You must do work to move any part of a circuit in a magnetic field. The work that you do is used to make the induced current flow in the presence of the circuit's resistance.

Generators as Energy Converters

Example 29.6 shows that the slidewire generator doesn't produce electrical energy out of nowhere; the energy is supplied by whatever object exerts the force that keeps the rod moving. All that the generator does is *convert* that energy into a different form. The equality between the rate at which *mechanical* energy is supplied to a generator and the rate at which *electrical* energy is generated holds for all types of generators, including the alternator described in Example 29.3. (We are ignoring the effects of friction in the bearings of an alternator or between the rod and the U-shaped conductor of a slidewire generator. The energy lost to friction is not available for conversion to electrical energy, so in real generators the friction is kept to a minimum.)

In Chapter 27 we stated that the magnetic force on moving charges can never do work. You might think, however, that the magnetic force $\vec{F} = I\vec{L} \times \vec{B}$ in Example 29.6 *is* doing (negative) work on the current-carrying rod as it moves, contradicting our earlier statement. In fact, the work done by the magnetic force is zero. The moving charges that make up the current in the rod in Fig. 29.12 have a vertical component of velocity, causing a horizontal component of force on these charges. As a result, there is a horizontal displacement of charge within the rod, the left side acquiring a net positive charge and the right side a net negative charge. The result is a horizontal component of electric field, perpendicular to the length of the rod (analogous to the Hall effect, described in Section 27.9). It is this field, in the direction of motion of the rod, that does work on the mobile charges in the rod and hence indirectly on the atoms making up the rod.



TEST YOUR UNDERSTANDING OF SECTION 29.2 The accompanying figure shows a wire coil being squeezed in a uniform magnetic field. (a) While the coil is being squeezed, is the induced emf in the coil (i) clockwise, (ii) counterclockwise, or (iii) zero? (b) Once the coil has reached its final squeezed shape, is the induced emf in the coil (i) clockwise, (ii) counterclockwise, or (iii) zero?

ANSWER

(a) **(i)**, (b) **(iii)** In (a), initially there is magnetic flux into the plane of the page, which we call positive. While the loop is being squeezed, the flux is becoming less positive ($d\Phi_B/dt < 0$) and so the induced emf is positive as in Fig. 29.6b ($\mathcal{E} = -d\Phi_B/dt > 0$). If you point the thumb of your right hand into the page, your fingers curl clockwise, so this is the direction of positive induced emf. In (b), since the coil's shape is no longer changing, the magnetic flux is not changing and there is no induced emf.

29.3 LENZ'S LAW

Lenz's law is a convenient alternative method for determining the direction of an induced current or emf. Lenz's law, named for the Russian physicist H. F. E. Lenz (1804–1865), is not an independent principle; it can be derived from Faraday's law. It always gives the same results as the sign rules we introduced in connection with Faraday's law, but it is often easier to use. Lenz's law also helps us gain intuitive understanding of various induction effects and of the role of energy conservation.

LENZ'S LAW The direction of any magnetic induction effect is such as to oppose the cause of the effect.

The "cause" may be changing flux through a stationary circuit due to a varying magnetic field, changing flux due to motion of the conductors that make up the circuit, or any combination. If the flux in a stationary circuit changes, as in Examples 29.1 and 29.2, the induced current sets up a magnetic field of its own. Within the area bounded by the circuit, this field is *opposite* to the original field if the original field is *increasing* but is in the *same* direction as the original field if the latter is *decreasing*. That is, the induced current opposes the *change in flux* through the circuit (*not* the flux itself).

If the flux change is due to motion of the conductors, as in Examples 29.3 through 29.6, the direction of the induced current in the moving conductor is such that the direction of the magnetic-field force on the conductor is opposite in direction to its motion. Thus the motion of the conductor, which caused the induced current, is opposed. We saw this explicitly for the slidewire generator in Example 29.6. In all these cases the induced current tries to preserve the *status quo* by opposing motion or a change of flux.

Lenz's law is also directly related to energy conservation. If the induced current in Example 29.6 were in the direction opposite to that given by Lenz's law, the magnetic force on the rod would accelerate it to ever-increasing speed with no external energy source, even though electrical energy is being dissipated in the circuit. This would be a clear violation of energy conservation and doesn't happen in nature.

CONCEPTUAL EXAMPLE 29.7 Lenz's law and the slidewire generator

WITH VARIATION PROBLEMS

In Fig. 29.11, the induced current in the loop causes an additional magnetic field in the area bounded by the loop. The direction of the induced current is counterclockwise, so from the discussion of Section 28.5, this additional magnetic field is directed *out of* the plane of the figure. That direction is opposite that of the original magnetic field, so it tends to cancel the effect of that field. This is just what Lenz's law predicts.

KEYCONCEPT A change in the magnetic flux through a circuit induces a current that produces an additional magnetic field of its own. This induced field always opposes the change in the magnetic flux (Lenz's law).

CONCEPTUAL EXAMPLE 29.8 Lenz's law and the direction of induced current

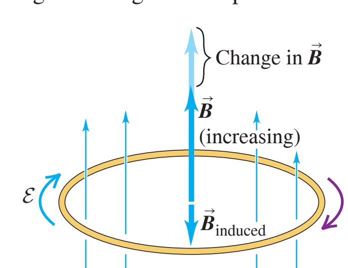
WITH VARIATION PROBLEMS

In **Fig. 29.13** there is a uniform magnetic field $\vec{B}$ through the coil. The magnitude of the field is increasing, so there is an induced emf. Use Lenz's law to determine the direction of the resulting induced current.

SOLUTION This situation is the same as in Example 29.1 (Section 29.2). By Lenz's law the induced current must produce a magnetic field $\vec{B}_{\text{induced}}$ inside the coil that is downward, opposing the change in flux. From the right-hand rule we described in Section 28.5 for the direction of the magnetic field produced by a circular loop, $\vec{B}_{\text{induced}}$ will be in the desired direction if the induced current flows as shown in Fig. 29.13.

Figure 29.14 (next page) shows several applications of Lenz's law to the similar situation of a magnet moving near a conducting loop. In each case, the induced current produces a magnetic field whose direction opposes the change in flux through the loop due to the magnet's motion.

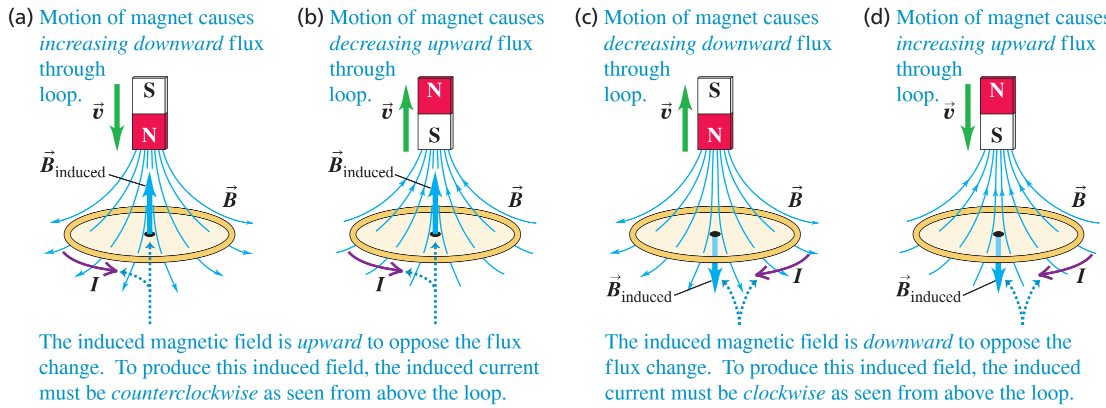
Figure 29.13 The induced current due to the change in $\vec{B}$ is clockwise, as seen from above the loop. The added field $\vec{B}_{\text{induced}}$ that it causes is downward, opposing the change in the upward field $\vec{B}$.



KEYCONCEPT Whenever there is a changing magnetic flux through a circuit, you can use Lenz's law to decide what the induced current direction will be: The current must produce an induced magnetic field that opposes the change in the flux.

Continued

Figure 29.14 Directions of induced currents as a bar magnet moves along the axis of a conducting loop. If the bar magnet is stationary, there is no induced current.



Lenz's Law and the Response to Flux Changes

Since an induced current always opposes any change in magnetic flux through a circuit, how is it possible for the flux to change at all? The answer is that Lenz's law gives only the *direction* of an induced current; the *magnitude* of the current depends on the resistance of the circuit. The greater the circuit resistance, the less the induced current that appears to oppose any change in flux and the easier it is for a flux change to take effect. If the loop in Fig. 29.14 were made out of wood (an insulator), there would be almost no induced current in response to changes in the flux through the loop.

Conversely, the less the circuit resistance, the greater the induced current and the more difficult it is to change the flux through the circuit. If the loop in Fig. 29.14 is a good conductor, an induced current flows as long as the magnet moves relative to the loop. Once the magnet and loop are no longer in relative motion, the induced current very quickly decreases to zero because of the nonzero resistance in the loop.

The extreme case occurs when the resistance of the circuit is *zero*. Then the induced current in Fig. 29.14 will continue to flow even after the induced emf has disappeared—that is, even after the magnet has stopped moving relative to the loop. Thanks to this *persistent current,* it turns out that the flux through the loop is exactly the same as it was before the magnet started to move, so the flux through a loop of zero resistance *never* changes. Exotic materials called *superconductors* do indeed have zero resistance; we discuss these further in Section 29.8.

TEST YOUR UNDERSTANDING OF SECTION 29.3 (a) Suppose the magnet in Fig. 29.14a were stationary and the loop of wire moved upward. Would the induced current in the loop be (i) in the same direction as shown in Fig. 29.14a, (ii) in the direction opposite to that shown in Fig. 29.14a, or (iii) zero? (b) Suppose the magnet and loop of wire in Fig. 29.14a both moved downward at the same velocity. Would the induced current in the loop be (i) in the same direction as shown in Fig. 29.14a, (ii) in the direction opposite to that shown in Fig. 29.14a, or (iii) zero?

ANSWER

(a) **(i)**, (b) **(iii)** In (a), as in the original situation, the magnet and loop are approaching each other and the downward flux through the loop is increasing. Hence the induced emf and induced current are the same. In (b), since the magnet and loop are moving together, the flux through the loop is not changing and no emf is induced.

29.4 MOTIONAL EMF

We've seen several situations in which a conductor moves in a magnetic field, as in the generators discussed in Examples 29.3 through 29.6. We can gain additional insight into the origin of the induced emf in these situations by considering the magnetic forces on mobile charges in the conductor. **Figure 29.15a** shows the same moving rod that we discussed in Example 29.5, separated for the moment from the U-shaped conductor. The magnetic field $\vec{B}$ is uniform and directed into the page, and we move the rod to the right at a constant velocity $\vec{v}$. A charged particle q in the rod then experiences a magnetic force $\vec{F} = q\vec{v} \times \vec{B}$ with magnitude $F = |q|vB$. We'll assume in the following discussion that q is positive; in that case the direction of this force is upward along the rod, from b toward a.

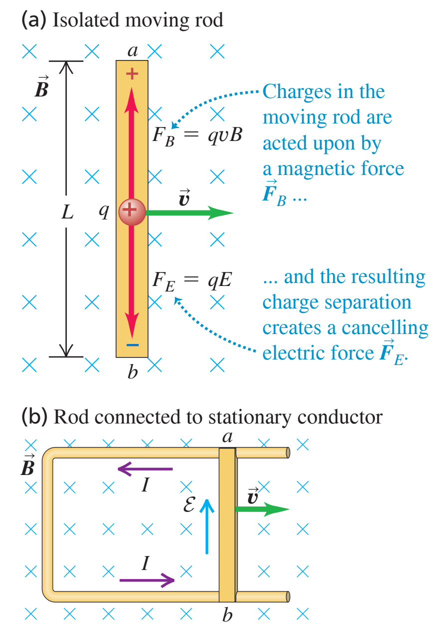
This magnetic force causes the free charges in the rod to move, creating an excess of positive charge at the upper end a and negative charge at the lower end b. This in turn creates an electric field $\vec{E}$ within the rod, in the direction from a toward b (opposite to the magnetic force). Charge continues to accumulate at the ends of the rod until $\vec{E}$ becomes large enough for the downward electric force (with magnitude qE) to cancel exactly the *upward* magnetic force (with magnitude qvB). Then $qE = qvB$ and the charges are in equilibrium.

The magnitude of the potential difference $V_{ab} = V_a - V_b$ is equal to the electric-field magnitude E multiplied by the length L of the rod. From the above discussion, $E = vB$, so

$$V_{ab} = EL = vBL \tag{29.5}$$

with point a at higher potential than point b.

Figure **29.15** A conducting rod moving in a uniform magnetic field. (a) The rod, the velocity, and the field are mutually perpendicular. (b) Direction of induced current in the circuit.



Now suppose the moving rod slides along a stationary U-shaped conductor, forming a complete circuit (Fig. 29.15b). No *magnetic* force acts on the charges in the stationary U-shaped conductor, but the charge that was near points a and b redistributes itself along the stationary conductor, creating an *electric* field within it. This field establishes a current in the direction shown. The moving rod has become a source of emf; within it, charge moves from lower to higher potential, and in the remainder of the circuit, charge moves from higher to lower potential. We call this emf a **motional emf,** denoted by $\mathcal{E}$. From the above discussion, the magnitude of this emf is

Motional emf, conductor length and velocity perpendicular to uniform $\vec{B}$

$$\mathcal{E} = vBL \tag{29.6}$$

Conductor speed (v); Conductor length (L); Magnitude of uniform magnetic field (B)

This corresponds to a force per unit charge of magnitude vB acting for a distance L along the moving rod. If the total circuit resistance of the U-shaped conductor and the sliding rod is R, the induced current I in the circuit is given by $vBL = IR$. This is the same result we obtained in Section 29.2 by using Faraday's law, and indeed motional emf is a particular case of Faraday's law. Verify that if we express v in meters per second, B in teslas, and L in meters, then $\mathcal{E}$ is in volts. (Recall that $1\ \text{V} = 1\ \text{J/C} = 1\ \text{T}\cdot\text{m}^2/\text{s}$.)

The emf associated with the moving rod in Fig. 29.15b is analogous to that of a battery with its positive terminal at a and its negative terminal at b, although the origins of the two emfs are quite different. In each case a nonelectrostatic force acts on the charges in the device, in the direction from b to a, and the emf is the work per unit charge done by this force when a charge moves from b to a in the device. When the device is connected to an external circuit, the direction of current is from b to a in the device and from a to b in the external circuit. Note that a motional emf is also present in the isolated moving rod in Fig. 29.15a, in the same way that a battery has an emf even when it's not part of a circuit.

You can determine the direction of the induced emf in Fig. 29.15 by using Lenz's law, even if (as in Fig. 29.15a) the conductor does not form a complete circuit. In this case we can mentally complete the circuit between the ends of the conductor and use Lenz's law to determine the direction of the current. From this we can deduce the polarity of the ends of the open-circuit conductor. The direction from the − end to the + end within the conductor is the direction the current would have if the circuit were complete.

Motional emf: General Form

We can generalize the concept of motional emf for a conductor with *any* shape, moving in any magnetic field, uniform or not, if we assume that the magnetic field at each point does not vary with time (**Fig. 29.16**). For an element $d\vec{l}$ of the conductor, the contribution $d\mathcal{E}$ to the emf is the magnitude dl multiplied by the component of $\vec{v} \times \vec{B}$ (the magnetic force per unit charge) parallel to $d\vec{l}$; that is,

$$d\mathcal{E} = (\vec{v} \times \vec{B}) \cdot d\vec{l}$$

For any closed conducting loop, the total emf is

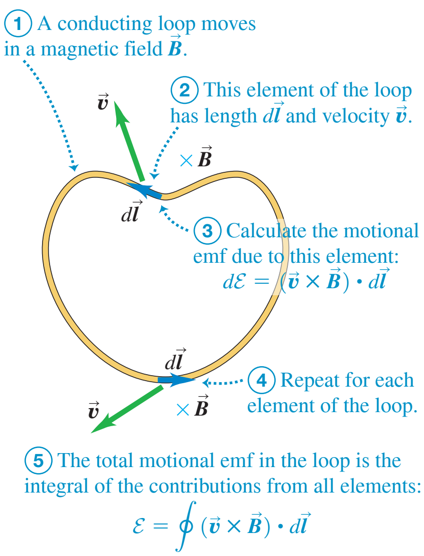
Line integral over all elements of **closed conducting loop**

Motional emf, general case

$$\mathcal{E} = \oint (\vec{v} \times \vec{B}) \cdot d\vec{l} \qquad (29.7)$$

Velocity of conductor element; Magnetic field at position of element; Length vector of conductor element

Figure 29.16 Calculating the motional emf for a moving current loop. The velocity $\vec{v}$ can be different for different elements if the loop is rotating or changing shape. The magnetic field $\vec{B}$ can also have different values at different points around the loop.



This expression looks very different from our original statement of Faraday's law, $\mathcal{E} = -d\Phi_B/dt$ [Eq. (29.3)]. In fact, though, the two statements are equivalent. It can be shown that the rate of change of magnetic flux through a moving conducting loop is always given by the negative of the expression in Eq. (29.7). Thus this equation gives us an alternative formulation of Faraday's law that is often convenient in problems with *moving* conductors. But when we have *stationary* conductors in changing magnetic fields, Eq. (29.7) *cannot* be used; in this case, $\mathcal{E} = -d\Phi_B/dt$ is the only correct way to express Faraday's law.

EXAMPLE 29.9 Motional emf in the slidewire generator

WITH VARIATION PROBLEMS

Suppose the moving rod in Fig. 29.15b is 0.10 m long, the velocity v is 2.5 m/s, the total resistance of the loop is 0.030 Ω, and B is 0.60 T. Find the motional emf, the induced current, and the force acting on the rod.

IDENTIFY and SET UP We'll find the motional emf $\mathcal{E}$ from Eq. (29.6) and the current from the values of $\mathcal{E}$ and the resistance R. The force on the rod is a *magnetic* force, exerted by $\vec{B}$ on the current in the rod; we'll find this force by using $\vec{F} = I\vec{L} \times \vec{B}$.

EXECUTE From Eq. (29.6) the motional emf is

$$\mathcal{E} = vBL = (2.5\ \text{m/s})(0.60\ \text{T})(0.10\ \text{m}) = 0.15\ \text{V}$$

The induced current in the loop is

$$I = \frac{\mathcal{E}}{R} = \frac{0.15\ \text{V}}{0.030\ \Omega} = 5.0\ \text{A}$$

In the expression for the magnetic force, $\vec{F} = I\vec{L} \times \vec{B}$, the vector $\vec{L}$ points in the same direction as the induced current in the rod (from b to a in Fig. 29.15). The right-hand rule for vector products shows that this force is directed *opposite* to the rod's motion. Since $\vec{L}$ and $\vec{B}$ are perpendicular, the force has magnitude

$$F = ILB = (5.0\ \text{A})(0.10\ \text{m})(0.60\ \text{T}) = 0.30\ \text{N}$$

EVALUATE We can check our answer for the direction of $\vec{F}$ by using Lenz's law. If we take the area vector $\vec{A}$ to point into the plane of the loop, the magnetic flux is positive and increasing as the rod moves to the right and increases the area of the loop. Lenz's law tells us that a force appears to oppose this increase in flux. Hence the force on the rod is to the left, opposite its motion.

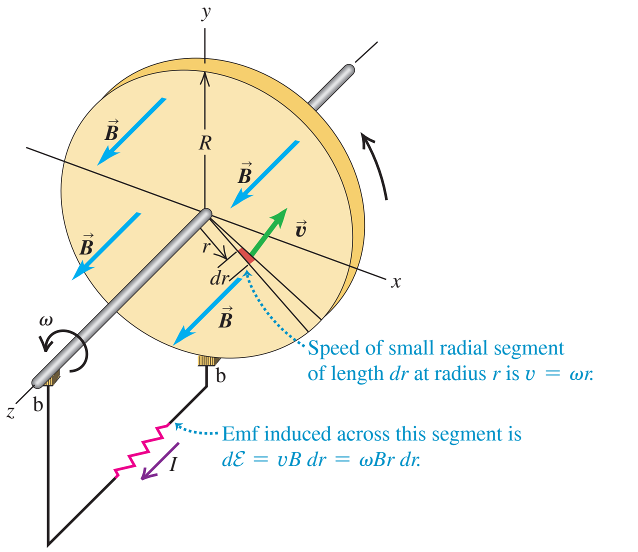
KEYCONCEPT When a straight conductor moves through a magnetic field, the magnetic forces on the mobile charges give rise to a motional emf. This emf is proportional to the field strength and to the speed and length of the conductor.

EXAMPLE 29.10 The Faraday disk dynamo

Figure 29.17 shows a conducting disk with radius R that lies in the xy-plane and rotates with constant angular velocity ω about the z-axis. The disk is in a uniform, constant $\vec{B}$ field in the z-direction. Find the induced emf between the center and the rim of the disk.

IDENTIFY and SET UP A motional emf arises because the conducting disk moves relative to $\vec{B}$. The complication is that different parts of the disk move at different speeds v, depending on their distance from the rotation axis. We'll address this by considering small segments of the disk and integrating their contributions to determine our target variable, the emf between the center and the rim. Consider the small segment of the disk shown in red in Fig. 29.17 and labeled by its velocity vector $\vec{v}$. The magnetic force per unit charge on this segment is $\vec{v} \times \vec{B}$, which points radially outward from the center of the disk. Hence the induced emf tends to make a current flow radially outward, which tells us that the moving conducting path to think about here is a straight line from the center to the rim. We can find the emf from each small disk segment along this line by using $d\mathcal{E} = (\vec{v} \times \vec{B}) \cdot d\vec{l}$ and then integrate to find the total emf.

Figure 29.17 A conducting disk with radius R rotating at an angular speed ω in a magnetic field $\vec{B}$. The emf is induced along radial lines of the disk and is applied to an external circuit through the two sliding contacts labeled b.



EXECUTE The length vector $d\vec{l}$ (of length dr) associated with the segment points radially outward, in the same direction as $\vec{v} \times \vec{B}$. The vectors $\vec{v}$ and $\vec{B}$ are perpendicular, and the magnitude of $\vec{v}$ is $v = \omega r$. The emf from the segment is then $d\mathcal{E} = \omega r B\, dr$. The total emf is the integral of $d\mathcal{E}$ from the center $(r = 0)$ to the rim $(r = R)$:

$$\mathcal{E} = \int_0^R \omega B r\, dr = \tfrac{1}{2}\omega B R^2$$

EVALUATE We can use this device as a source of emf in a circuit by completing the circuit through two stationary brushes (labeled b in the figure) that contact the disk and its conducting shaft as shown. Such a disk is called a *Faraday disk dynamo* or a *homopolar generator*. Unlike the alternator in Example 29.3, the Faraday disk dynamo is a direct-current generator; it produces an emf that is constant in time. Can you use Lenz's law to show that for the direction of rotation in Fig. 29.17, the current in the external circuit must be in the direction shown?

KEYCONCEPT In general, finding the induced emf in a conductor requires doing an integral. This is because the magnetic force on a mobile charge can be different at different locations in the conductor.

TEST YOUR UNDERSTANDING OF SECTION 29.4 The earth's magnetic field points toward (magnetic) north. For simplicity, assume that the field has no vertical component (as is the case near the earth's equator). (a) If you hold a metal rod in your hand and walk toward the east, how should you orient the rod to get the maximum motional emf between its ends? (i) East-west; (ii) north-south; (iii) up-down; (iv) you get the same motional emf with all of these orientations. (b) How should you hold it to get *zero* emf as you walk toward the east? (i) East-west; (ii) north-south; (iii) up-down; (iv) none of these. (c) In which direction should you travel so that the motional emf across the rod is zero no matter how the rod is oriented? (i) West; (ii) north; (iii) south; (iv) straight up; (v) straight down.

ANSWER

(a) **(iii)**; (b) **(i) or (ii)**; (c) **(ii) or (iii)** You'll get the maximum motional emf if you hold the rod vertically, so that its length is perpendicular to both the magnetic field and the direction of motion. With this orientation, $\vec{L}$ is parallel to $\vec{v} \times \vec{B}$. If you hold the rod in any horizontal orientation, $\vec{L}$ will be perpendicular to $\vec{v} \times \vec{B}$ and no emf will be induced. If you walk due north or south, $\vec{v} \times \vec{B} = \mathbf{0}$ and no emf will be induced for any orientation of the rod.

29.5 INDUCED ELECTRIC FIELDS

When a conductor moves in a magnetic field, we can understand the induced emf on the basis of magnetic forces on charges in the conductor, as described in Section 29.4. But an induced emf also occurs when there is a changing flux through a stationary conductor. What is it that pushes the charges around the circuit in this type of situation?

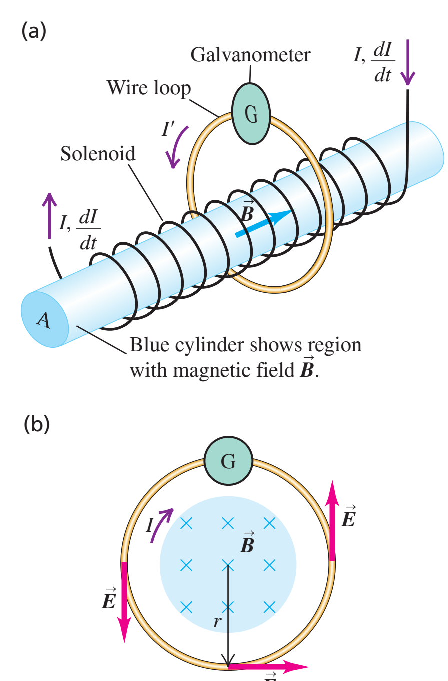
Let's consider the situation shown in **Fig. 29.18**. A long, thin solenoid with cross-sectional area A and n turns per unit length is encircled at its center by a circular conducting loop. The galvanometer G measures the current in the loop. A current I in the winding of the solenoid sets up a magnetic field $\vec{B}$ along the solenoid axis, as shown, with magnitude B as calculated in Section 28.7: $B = \mu_0 n I$, where n is the number of turns per unit length. If we ignore the small field outside the solenoid and take the area vector $\vec{A}$ to point in the same direction as $\vec{B}$, then the magnetic flux Φ_B through the loop is

$$\Phi_B = BA = \mu_0 n I A$$

When the solenoid current I changes with time, the magnetic flux Φ_B also changes, and according to Faraday's law the induced emf in the loop is given by

$$\mathcal{E} = -\frac{d\Phi_B}{dt} = -\mu_0 n A \frac{dI}{dt} \tag{29.8}$$

Figure 29.18 (a) The windings of a long solenoid carry a current I that is increasing at a rate dI/dt. The magnetic flux in the solenoid is increasing at a rate $d\Phi_B/dt$, and this changing flux passes through a wire loop. An emf $\mathcal{E} = -d\Phi_B/dt$ is induced in the loop, inducing a current I' that is measured by the galvanometer G. (b) Cross-sectional view.



If the total resistance of the loop is R, the induced current in the loop, which we may call I', is $I' = \mathcal{E}/R$.

But what *force* makes the charges move around the wire loop? It can't be a magnetic force because the loop isn't even *in* a magnetic field. We are forced to conclude that there has to be an **induced electric field** in the conductor *caused by the changing magnetic flux*. Induced electric fields are *very* different from the electric fields caused by charges, which we discussed in Chapter 23. To see this, note that when a charge q goes once around the loop, the total work done on it by the electric field must be equal to q times the emf $\mathcal{E}$. That is, the electric field in the loop *is not conservative,* as we used the term in Section 23.1, because the line integral of $\vec{E}$ around a closed path is not zero. Indeed, this line integral, representing the work done by the induced $\vec{E}$ field per unit charge, is equal to the induced emf $\mathcal{E}$:

$$\oint \vec{E} \cdot d\vec{l} = \mathcal{E} \tag{29.9}$$

From Faraday's law the emf $\mathcal{E}$ is also the negative of the rate of change of magnetic flux through the loop. Thus for this case we can restate Faraday's law as

Faraday's law for a stationary integration path:

$$\overbrace{\oint \vec{E} \cdot d\vec{l}}^{\text{Line integral of electric field around path}} = \underbrace{-\frac{d\Phi_B}{dt}}_{\text{Negative of the time rate of change of magnetic flux through path}} \tag{29.10}$$

Note that Faraday's law is *always* true in the form $\mathcal{E} = -d\Phi_B/dt$; the form given in Eq. (29.10) is valid *only* if the path around which we integrate is *stationary*.

Let's apply Eq. (29.10) to the stationary circular loop in Fig. 29.18b, which we take to have radius r. Because of cylindrical symmetry, $\vec{E}$ has the same magnitude at every point on the circle and is tangent to it at each point. (Symmetry would also permit the field to be *radial,* but then Gauss's law would require the presence of a net charge inside the circle, and there is none.) The line integral in Eq. (29.10) becomes simply the magnitude E times the circumference $2\pi r$ of the loop, $\oint \vec{E} \cdot d\vec{l} = 2\pi r E$, and Eq. (29.10) gives

$$E = \frac{1}{2\pi r}\left|\frac{d\Phi_B}{dt}\right| \tag{29.11}$$

The directions of $\vec{E}$ at points on the loop are shown in Fig. 29.18b. We know that $\vec{E}$ has to have the direction shown when $\vec{B}$ in the solenoid is increasing, because $\oint \vec{E} \cdot d\vec{l}$ has to be negative when $d\Phi_B/dt$ is positive. The same approach can be used to find the induced electric field *inside* the solenoid when the solenoid $\vec{B}$ field is changing; we leave the details to you.

Figure **29.19** Applications of induced electric fields. (a) This car is powered by an electric motor. As the car comes to a halt, the spinning wheels run the motor backward so that it acts as a generator. The resulting induced current is used to recharge the car's batteries. (b) The rotating crankshaft of a piston-engine airplane spins a magnet, inducing an emf in an adjacent coil and generating the spark that ignites fuel in the engine cylinders. This keeps the engine running even if the airplane's other electrical systems fail.

Nonelectrostatic Electric Fields

We've learned that Faraday's law, Eq. (29.3), is valid for two rather different situations. In one, an emf is induced by magnetic forces on charges when a conductor moves through a magnetic field. In the other, a time-varying magnetic field induces an electric field and hence an emf; the $\vec{E}$ field is induced even when no conductor is present. This $\vec{E}$ field differs from an electro*static* field in an important way. It is *nonconservative;* the line integral $\oint \vec{E} \cdot d\vec{l}$ around a closed path is not zero, and when a charge moves around a closed path, the field does a nonzero amount of work on it. It follows that for such a field the concept of *potential* has no meaning. We call such a field a **nonelectrostatic field.** In contrast, an *electrostatic* field is *always* conservative, as we discussed in Section 23.1, and always has an associated potential function. Despite this difference, the fundamental effect of *any* electric field is to exert a force $\vec{F} = q\vec{E}$ on a charge q. This relationship is valid whether $\vec{E}$ is conservative and produced by charges or nonconservative and produced by changing magnetic flux.

So a changing magnetic field acts as a source of electric field of a sort that we *cannot* produce with any static charge distribution. What's more, we'll see in Section 29.7 that a changing *electric* field acts as a source of *magnetic* field. We'll explore this symmetry between the two fields in our study of electromagnetic waves in Chapter 32.

If any doubt remains in your mind about the reality of magnetically induced electric fields, consider a few of the many practical applications (**Fig. 29.19**). Pickups in electric guitars use currents induced in stationary pickup coils by the vibration of nearby ferromagnetic strings. Alternators in most cars use rotating magnets to induce currents in stationary coils. Whether we realize it or not, magnetically induced electric fields play an important role in everyday life.

EXAMPLE 29.11 Induced electric fields

WITH VARIATION PROBLEMS

Suppose the long solenoid in Fig. 29.18a has 500 turns per meter and cross-sectional area 4.0 cm^2. The current in its windings is increasing at 100 A/s. (a) Find the magnitude of the induced emf in the wire loop outside the solenoid. (b) Find the magnitude of the induced electric field within the loop if its radius is 2.0 cm.

IDENTIFY and SET UP As in Fig. 29.18b, the increasing magnetic field inside the solenoid causes a change in the magnetic flux through the wire loop and hence induces an electric field $\vec{E}$ around the loop. Our target variables are the induced emf $\mathcal{E}$ and the electric-field magnitude E. We use Eq. (29.8) to determine the emf. The loop and the solenoid share the same central axis. Hence, by symmetry, the electric field is tangent to the loop and has the same magnitude E all the way around its circumference. We can therefore use Eq. (29.9) to find E.

EXECUTE (a) From Eq. (29.8), the induced emf is

$$\begin{aligned}\mathcal{E} &= -\frac{d\Phi_B}{dt} = -\mu_0 nA\frac{dI}{dt}\\ &= -(4\pi \times 10^{-7}\ \text{Wb/A}\cdot\text{m})(500\ \text{turns/m})(4.0 \times 10^{-4}\ \text{m}^2)(100\ \text{A/s})\\ &= -25 \times 10^{-6}\ \text{Wb/s} = -25 \times 10^{-6}\ \text{V} = -25\ \mu\text{V}\end{aligned}$$

(b) By symmetry the line integral $\oint \vec{E}\cdot d\vec{l}$ has absolute value $2\pi rE$ no matter which direction we integrate around the loop. This is equal to the absolute value of the emf, so

$$E = \frac{|\mathcal{E}|}{2\pi r} = \frac{25 \times 10^{-6}\ \text{V}}{2\pi(2.0 \times 10^{-2}\ \text{m})} = 2.0 \times 10^{-4}\ \text{V/m}$$

EVALUATE In Fig. 29.18b the magnetic flux *into* the plane of the figure is increasing. According to the right-hand rule for induced emf (Fig. 29.6), a positive emf would be clockwise around the loop; the negative sign of $\mathcal{E}$ shows that the emf is in the counterclockwise direction. Can you also show this by using Lenz's law?

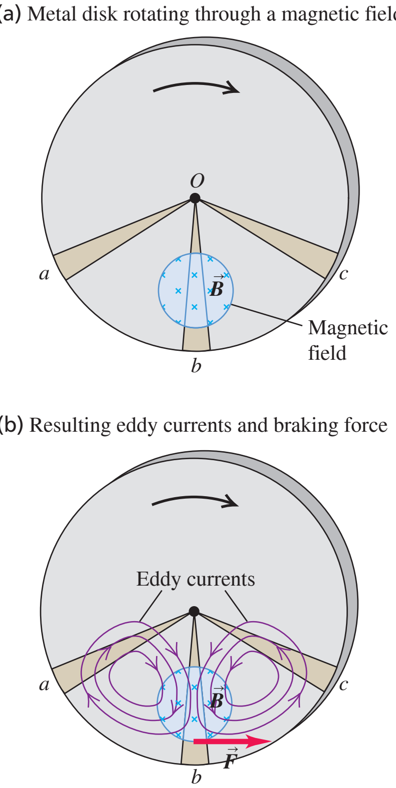
KEYCONCEPT A changing magnetic flux through a closed loop produces an induced electric field $\vec{E}$. The line integral of $\vec{E}$ around the loop is equal to the induced emf in the loop.

TEST YOUR UNDERSTANDING OF SECTION 29.5 If you wiggle a magnet back and forth in your hand, are you generating an electric field? If so, is this field conservative?

ANSWER

yes, no The magnetic field at a fixed position changes as you move the magnet, which induces an electric field. Such induced electric fields are *not* conservative.

Figure 29.20 Eddy currents induced in a rotating metal disk.



(a) Metal disk rotating through a magnetic field

(b) Resulting eddy currents and braking force

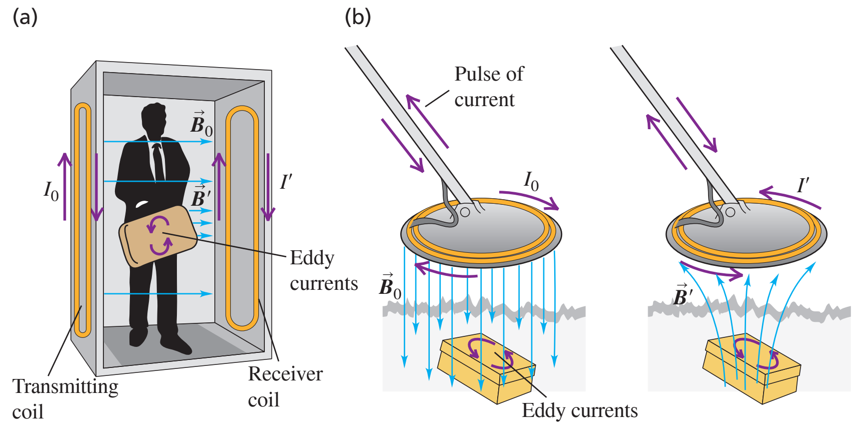
29.6 EDDY CURRENTS

In the examples of induction effects that we have studied, the induced currents have been confined to well-defined paths in conductors and other components forming a circuit. However, many pieces of electrical equipment contain masses of metal moving in magnetic fields or located in changing magnetic fields. In situations like these we can have induced currents that circulate throughout the volume of a material. Because their flow patterns resemble swirling eddies in a river, we call these **eddy currents.**

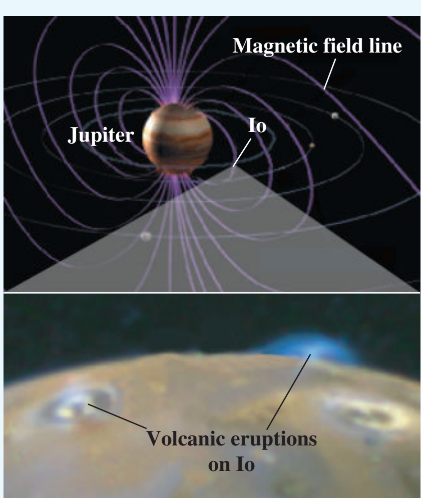
As an example, consider a metallic disk rotating in a magnetic field perpendicular to the plane of the disk but confined to a limited portion of the disk's area, as shown in **Fig. 29.20a**. Sector *Ob* is moving across the field and has an emf induced in it. Sectors *Oa* and *Oc* are not in the field, but they provide return conducting paths for charges displaced along *Ob* to return from *b* to *O*. The result is a circulation of eddy currents in the disk, somewhat as sketched in Fig. 29.20b.

We can use Lenz's law to decide on the direction of the induced current in the neighborhood of sector *Ob*. This current must experience a magnetic force $\vec{F} = I\vec{L} \times \vec{B}$ that *opposes* the rotation of the disk, and so this force must be to the right in Fig. 29.20b. Since $\vec{B}$ is directed into the plane of the disk, the current and hence $\vec{L}$ have downward components. The return currents lie outside the field, so they do not experience magnetic forces. The interaction between the eddy currents and the field causes a braking action on the disk. Such effects can be used to stop the rotation of a circular saw quickly when

Figure **29.21** (a) A metal detector at an airport security checkpoint generates an alternating magnetic field $\vec{B}_0$. This induces eddy currents in a conducting object carried through the detector. The eddy currents in turn produce an alternating magnetic field $\vec{B}'$, which induces a current in the detector's receiver coil. (b) Portable metal detectors work on the same principle.



APPLICATION Eddy Currents Help Power Io's Volcanoes Jupiter's moon Io is slightly larger than the earth's moon. It moves at more than 60,000 km/h through Jupiter's intense magnetic field (about ten times stronger than the earth's field), which sets up strong eddy currents within Io that dissipate energy at a rate of 10^{12} W. This dissipated energy helps to heat Io's interior and causes volcanic eruptions on its surface, as shown in the lower close-up image. (Gravitational effects from Jupiter cause even more heating.)



the power is turned off. Eddy current braking is used on some electrically powered rapid-transit vehicles. Electromagnets mounted in the cars induce eddy currents in the rails; the resulting magnetic fields cause braking forces on the electromagnets and thus on the cars.

Eddy currents have many other practical uses. In induction furnaces, eddy currents are used to heat materials in completely sealed containers for processes in which it is essential to avoid the slightest contamination of the materials. The metal detectors used at airport security checkpoints (**Fig. 29.21a**) operate by detecting eddy currents induced in metallic objects. Similar devices (Fig. 29.21b) are used to find buried treasure such as bottlecaps and lost pennies.

Eddy currents also have undesirable effects. In an alternating-current transformer, coils wrapped around an iron core carry a sinusoidally varying current. The resulting eddy currents in the core waste energy through I^2R heating and set up an unwanted opposing emf in the coils. To minimize these effects, the core is designed so that the paths for eddy currents are as narrow as possible. We'll describe how this is done when we discuss transformers in Section 31.6.

TEST YOUR UNDERSTANDING OF SECTION 29.6 Suppose that the magnetic field in Fig. 29.20 were directed out of the plane of the figure and the disk were rotating counterclockwise. Compared to the directions of the force $\vec{F}$ and the eddy currents shown in Fig. 29.20b, what would the new directions be? (i) The force $\vec{F}$ and the eddy currents would both be in the same direction; (ii) the force $\vec{F}$ would be in the same direction, but the eddy currents would be in the opposite direction; (iii) the force $\vec{F}$ would be in the opposite direction, but the eddy currents would be in the same direction; (iv) the force $\vec{F}$ and the eddy currents would be in the opposite directions.

ANSWER

(iii) By Lenz's law, the force must oppose the motion of the disk through the magnetic field. Since the disk material is now moving to the right through the field region, the force $\vec{F}$ is to the left—that is, in the opposite direction to that shown in Fig. 29.20b. To produce a leftward magnetic force $\vec{F} = I\vec{L} \times \vec{B}$ on currents moving through a magnetic field $\vec{B}$ directed out of the plane of the figure, the eddy currents must be moving downward in the figure—that is, in the same direction shown in Fig. 29.20b.

29.7 DISPLACEMENT CURRENT AND MAXWELL'S EQUATIONS

We have seen that a varying magnetic field gives rise to an induced electric field. In one of the more remarkable examples of the symmetry of nature, it turns out that a varying *electric* field gives rise to a *magnetic* field. This effect is of tremendous importance, for it turns out to explain the existence of radio waves, gamma rays, and visible light, as well as all other forms of electromagnetic waves.

Generalizing Ampere's Law

To see the origin of the relationship between varying electric fields and magnetic fields, let's return to Ampere's law as given in Section 28.6, Eq. (28.20):

$$\oint \vec{B} \cdot d\vec{l} = \mu_0 I_{\text{encl}}$$

The problem with Ampere's law in this form is that it is *incomplete*. To see why, let's consider the process of charging a capacitor (**Fig. 29.22**). Conducting wires lead current i_C into one plate and out of the other; the charge Q increases, and the electric field $\vec{E}$ between the plates increases. The notation i_C indicates *conduction* current to distinguish it from another kind of current we are about to encounter, called *displacement* current i_D. We use lowercase i's and v's to denote instantaneous values of currents and potential differences, respectively, that may vary with time.

Let's apply Ampere's law to the circular path shown. The integral $\oint \vec{B} \cdot d\vec{l}$ around this path equals $\mu_0 I_{\text{encl}}$. For the plane circular area bounded by the circle, I_{encl} is just the current i_C in the left conductor. But the surface that bulges out to the right is bounded by the same circle, and the current through that surface is zero. So $\oint \vec{B} \cdot d\vec{l}$ is equal to $\mu_0 i_C$, and at the same time it is equal to zero! This is a clear contradiction.

However, something else is happening on the bulged-out surface. As the capacitor charges, the electric field $\vec{E}$ and the electric *flux* Φ_E through the surface are increasing. We can determine their rates of change in terms of the charge and current. The instantaneous charge is $q = Cv$, where C is the capacitance and v is the instantaneous potential difference. For a parallel-plate capacitor, $C = \epsilon_0 A/d$, where A is the plate area and d is the spacing. The potential difference v between plates is $v = Ed$, where E is the electric-field magnitude between plates. (We ignore fringing and assume that $\vec{E}$ is uniform in the region between the plates.) If this region is filled with a material with permittivity ϵ, we replace ϵ_0 by ϵ everywhere; we'll use ϵ in the following discussion.

Substituting these expressions for C and v into $q = Cv$, we can express the capacitor charge q in terms of the electric flux $\Phi_E = EA$ through the surface:

$$q = Cv = \frac{\epsilon A}{d}(Ed) = \epsilon EA = \epsilon \Phi_E \qquad (29.12)$$

As the capacitor charges, the rate of change of q is the conduction current, $i_C = dq/dt$. Taking the derivative of Eq. (29.12) with respect to time, we get

$$i_C = \frac{dq}{dt} = \epsilon \frac{d\Phi_E}{dt} \qquad (29.13)$$

Stretching our imagination a bit, we invent a fictitious **displacement current** i_D in the region between the plates, defined as

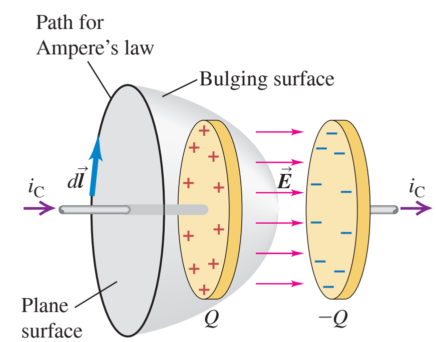
$$\underset{\text{through an area}}{\text{Displacement current}}\; i_D = \underset{\text{Permittivity of material in area}}{\epsilon}\; \underset{\text{Time rate of change of electric flux through area}}{\frac{d\Phi_E}{dt}} \qquad (29.14)$$

That is, we imagine that the changing flux through the curved (bulged-out) surface in Fig. 29.22 is equivalent, in Ampere's law, to a conduction current through that surface. We include this fictitious current, along with the real conduction current i_C, in Ampere's law:

$$\oint \vec{B} \cdot d\vec{l} = \mu_0 (i_C + i_D)_{\text{encl}} \quad \text{(generalized Ampere's law)} \qquad (29.15)$$

Ampere's law in this form is obeyed no matter which surface we use in Fig. 29.22. For the flat surface, i_D is zero; for the curved surface, i_C is zero; and i_C for the flat surface equals i_D for the curved surface. Equation (29.15) remains valid in a magnetic material, provided that the magnetization is proportional to the external field and we replace μ_0 by μ.

Figure **29.22** Parallel-plate capacitor being charged. The conduction current through the plane surface is i_C, but there is no conduction current through the surface that bulges out to pass between the plates. The two surfaces have a common boundary, so this difference in I_{encl} leads to an apparent contradiction in applying Ampere's law.



The fictitious displacement current i_D was invented in 1865 by the Scottish physicist James Clerk Maxwell. There is a corresponding *displacement current density* $j_D = i_D/A$; using $\Phi_E = EA$ and dividing Eq. (29.14) by A, we find

$$j_D = \epsilon \frac{dE}{dt} \tag{29.16}$$

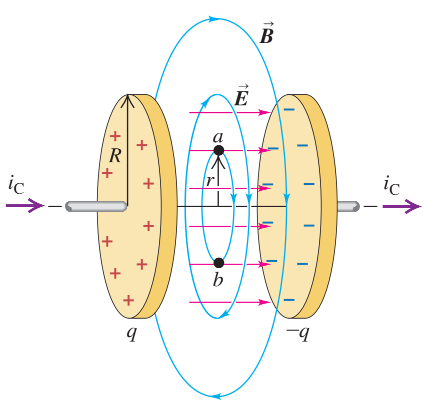
We have pulled the concept out of thin air, as Maxwell did, but we see that it enables us to save Ampere's law in situations such as that in Fig. 29.22.

Another benefit of displacement current is that it lets us generalize Kirchhoff's junction rule, discussed in Section 26.2. Considering the left plate of the capacitor, we have conduction current into it but none out of it. But when we include the displacement current, we have conduction current coming in one side and an equal displacement current coming out the other side. With this generalized meaning of the term "current," we can speak of current going *through* the capacitor.

The Reality of Displacement Current

You might well ask at this point whether displacement current has any real physical significance or whether it is just a ruse to satisfy Ampere's law and Kirchhoff's junction rule. Here's a fundamental experiment that helps to answer that question. We take a plane circular area between the capacitor plates (**Fig. 29.23**). If displacement current really plays the role in Ampere's law that we have claimed, then there ought to be a magnetic field in the region between the plates while the capacitor is charging. We can use our generalized Ampere's law, including displacement current, to predict what this field should be.

Figure 29.23 A capacitor being charged by a current i_C has a displacement current equal to i_C between the plates, with displacement-current density $j_D = \epsilon\, dE/dt$. This can be regarded as the source of the magnetic field between the plates.



To be specific, let's picture round capacitor plates with radius R. To find the magnetic field at a point in the region between the plates at a distance r from the axis, we apply Ampere's law to a circle of radius r passing through the point, with $r < R$. This circle passes through points a and b in Fig. 29.23. The total current enclosed by the circle is j_D times its area, or $(i_D/\pi R^2)(\pi r^2)$. The integral $\oint \vec{B} \cdot d\vec{l}$ in Ampere's law is just B times the circumference $2\pi r$ of the circle, and because $i_D = i_C$ for the charging capacitor, Ampere's law becomes

$$\oint \vec{B} \cdot d\vec{l} = 2\pi r B = \mu_0 \frac{r^2}{R^2} i_C \qquad \text{or}$$

$$B = \frac{\mu_0}{2\pi} \frac{r}{R^2} i_C \tag{29.17}$$

This result predicts that in the region between the plates $\vec{B}$ is zero at the axis and increases linearly with distance from the axis. A similar calculation shows that *outside* the region between the plates (that is, for $r > R$), $\vec{B}$ is the same as though the wire were continuous and the plates not present at all.

When we *measure* the magnetic field in this region, we find that it really is there and that it behaves just as Eq. (29.17) predicts. This confirms directly the role of displacement current as a source of magnetic field. It is now established beyond reasonable doubt that Maxwell's displacement current, far from being just an artifice, is a fundamental fact of nature.

Maxwell's Equations of Electromagnetism

We are now in a position to wrap up in a single package *all* of the relationships between electric and magnetic fields and their sources. This package consists of four equations, called **Maxwell's equations.** Maxwell did not discover all of these equations singlehandedly (though he did develop the concept of displacement current). But he did put them together and recognized their significance, particularly in predicting the existence of electromagnetic waves.

For now we'll state Maxwell's equations in their simplest form, for the case in which we have charges and currents in otherwise empty space. In Chapter 32 we'll discuss how to modify these equations if a dielectric or a magnetic material is present.

Two of Maxwell's equations involve an integral of $\vec{E}$ or $\vec{B}$ over a closed surface. The first is simply Gauss's law for electric fields, Eq. (22.8):

Gauss's law for $\vec{E}$:

Flux of electric field through a closed surface

Charge enclosed by surface

Electric constant

$$\overbrace{\oint \vec{E}\cdot d\vec{A}} = \frac{Q_{\text{encl}}}{\epsilon_0} \qquad (29.18)$$

The second is the analogous relationship for *magnetic* fields, Eq. (27.8):

Gauss's law for $\vec{B}$:

Flux of magnetic field through any closed surface ...

$$\overbrace{\oint \vec{B}\cdot d\vec{A}} = 0 \qquad (29.19)$$

... equals zero.

This statement means, among other things, that there are no magnetic monopoles (single magnetic charges) to act as sources of magnetic field.

The third and fourth equations involve a line integral of $\vec{E}$ or $\vec{B}$ around a closed path. Faraday's law states that a changing magnetic flux acts as a source of electric field:

Faraday's law for a stationary integration path:

Line integral of electric field around path

Negative of the time rate of change of magnetic flux through path

$$\overbrace{\oint \vec{E}\cdot d\vec{l}} = \underbrace{-\frac{d\Phi_B}{dt}} \qquad (29.20)$$

If there is a changing magnetic field, the line integral in Eq. (29.20)—which must be carried out over a *stationary* closed path—is not zero. Thus the $\vec{E}$ field produced by a changing $\vec{B}$ is not conservative.

The fourth and final equation is Ampere's law including displacement current. It states that both a conduction current and a changing electric flux act as sources of magnetic field:

Ampere's law for a stationary integration path:

Line integral of magnetic field around path

Electric constant

Time rate of change of electric flux through path

$$\overbrace{\oint \vec{B}\cdot d\vec{l}} = \mu_0\left(i_C + \underbrace{\epsilon_0 \overbrace{\frac{d\Phi_E}{dt}}}\right)_{\text{encl}} \qquad (29.21)$$

Magnetic constant

Conduction current through path

Displacement current through path

It's worthwhile to look more carefully at the electric field $\vec{E}$ and its role in Maxwell's equations. In general, the total $\vec{E}$ field at a point in space can be the superposition of an electrostatic field $\vec{E}_c$ caused by a distribution of charges at rest and a magnetically induced, nonelectrostatic field $\vec{E}_n$. That is,

$$\vec{E} = \vec{E}_c + \vec{E}_n$$

The electrostatic part $\vec{E}_c$ is *always* conservative, so $\oint \vec{E}_c \cdot d\vec{l} = 0$. This conservative part of the field does not contribute to the integral in Faraday's law, so we can take $\vec{E}$ in Eq. (29.20) to be the *total* electric field $\vec{E}$, including both the part $\vec{E}_c$ due to charges and the magnetically induced part $\vec{E}_n$. Similarly, the nonconservative part $\vec{E}_n$ of the $\vec{E}$ field does not contribute to the integral in Gauss's law, because this part of the field is not caused by static charges. Hence $\oint \vec{E}_n \cdot d\vec{A}$ is always zero. We conclude that in all the Maxwell equations, $\vec{E}$ is the total electric field; these equations don't distinguish between conservative and nonconservative fields.

Symmetry in Maxwell's Equations

There is a remarkable symmetry in Maxwell's four equations. In empty space where there is no charge, the first two equations [Eqs. (29.18) and (29.19)] are identical in form, one containing $\vec{E}$ and the other containing $\vec{B}$ (**Fig. 29.24**). When we compare the second two equations, Eq. (29.20) says that a changing magnetic flux creates an electric field, and Eq. (29.21) says that a changing electric flux creates a magnetic field. In empty space, where there is no conduction current, $i_C = 0$ and the two equations have the same form, apart from a numerical constant and a negative sign, with the roles of $\vec{E}$ and $\vec{B}$ exchanged.

Figure **29.24** Maxwell's equations in empty space are highly symmetric.

In empty space there are no charges, so the fluxes of $\vec{E}$ and $\vec{B}$ through any closed surface are equal to zero.

$$\oint \vec{E} \cdot d\vec{A} = 0$$

$$\oint \vec{B} \cdot d\vec{A} = 0$$

$$\oint \vec{E} \cdot d\vec{l} = -\frac{d\Phi_B}{dt}$$

$$\oint \vec{B} \cdot d\vec{l} = \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}$$

In empty space there are no conduction currents, so the line integrals of $\vec{E}$ and $\vec{B}$ around any closed path are related to the rate of change of flux of the other field.

We can rewrite Eqs. (29.20) and (29.21) in a different but equivalent form by introducing the definitions of magnetic flux, $\Phi_B = \int \vec{B} \cdot d\vec{A}$, and electric flux, $\Phi_E = \int \vec{E} \cdot d\vec{A}$, respectively. In empty space, where there is no charge or conduction current, $i_C = 0$ and $Q_{encl} = 0$, and we have

$$\oint \vec{E} \cdot d\vec{l} = -\frac{d}{dt}\int \vec{B} \cdot d\vec{A} \qquad (29.22)$$

$$\oint \vec{B} \cdot d\vec{l} = \mu_0 \epsilon_0 \frac{d}{dt}\int \vec{E} \cdot d\vec{A} \qquad (29.23)$$

Again we notice the symmetry between the roles of $\vec{E}$ and $\vec{B}$ in these expressions.

The most remarkable feature of these equations is that a time-varying field of *either* kind induces a field of the other kind in neighboring regions of space. Maxwell recognized that these relationships predict the existence of electromagnetic disturbances consisting of time-varying electric and magnetic fields that travel or *propagate* from one region of space to another, even if no matter is present in the intervening space. Such disturbances, called *electromagnetic waves,* provide the physical basis for light, radio and television waves, infrared, ultraviolet, and x rays. We'll return to this vitally important topic in Chapter 32.

Although it may not be obvious, *all* the basic relationships between fields and their sources are contained in Maxwell's equations. We can derive Coulomb's law from Gauss's law, we can derive the law of Biot and Savart from Ampere's law, and so on. When we add the equation that defines the $\vec{E}$ and $\vec{B}$ fields in terms of the forces that they exert on a charge q, namely,

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}) \qquad (29.24)$$

we have *all* the fundamental relationships of electromagnetism!

Maxwell's equations would have even greater symmetry between the $\vec{E}$ and $\vec{B}$ fields if single magnetic charges (magnetic monopoles) existed. The right side of Eq. (29.19) would contain the total *magnetic* charge enclosed by the surface, and the right side of Eq. (29.20) would include a magnetic monopole current term. However, no magnetic monopoles have yet been found.

In conciseness and generality, Maxwell's equations are in the same league with Newton's laws of motion and the laws of thermodynamics. Indeed, a major goal of science is learning how to express very broad and general relationships in a concise and compact form. Maxwell's synthesis of electromagnetism stands as a towering intellectual achievement, comparable to the Newtonian synthesis we described at the end of Section 13.5 and to the development of relativity and quantum mechanics in the 20th century.

TEST YOUR UNDERSTANDING OF SECTION 29.7 (a) Which of Maxwell's equations explains how a credit card reader works? (b) Which one describes how a wire carrying a steady current generates a magnetic field?

ANSWER

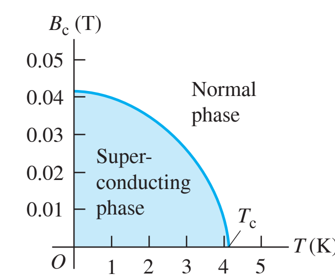
(a) **Faraday's law,** (b) **Ampere's law** A credit card reader works by inducing currents in the reader's coils as the card's magnetized stripe is swiped (see the answer to the chapter opening question). Ampere's law describes how currents of all kinds (both conduction currents and displacement currents) give rise to magnetic fields.

29.8 SUPERCONDUCTIVITY

The most familiar property of a superconductor is the sudden disappearance of all electrical resistance when the material is cooled below a temperature called the *critical temperature,* denoted by T_c. We discussed this behavior and the circumstances of its discovery in Section 25.2. But superconductivity is far more than just the absence of measurable resistance. As we'll see in this section, superconductors also have extraordinary *magnetic* properties.

The first hint of unusual magnetic properties was the discovery that for any superconducting material the critical temperature T_c changes when the material is placed in an externally produced magnetic field $\vec{B}_0$. **Figure 29.25** shows this dependence for mercury, the first element in which superconductivity was observed. As the external field magnitude B_0 increases, the superconducting transition occurs at lower and lower temperature. When B_0 is greater than 0.0412 T, *no* superconducting transition occurs. The minimum magnitude of magnetic field that is needed to eliminate superconductivity at a temperature below T_c is called the *critical field,* denoted by B_c.

Figure 29.25 Phase diagram for pure mercury, showing the critical magnetic field B_c and its dependence on temperature. Superconductivity is impossible above the critical temperature T_c. The curves for other superconducting materials are similar but with different numerical values.



The Meissner Effect

Another aspect of the magnetic behavior of superconductors appears if we place a homogeneous sphere of a superconducting material in a uniform applied magnetic field $\vec{B}_0$ at a temperature T greater than T_c. The material is then in the normal phase, not the superconducting phase (**Fig. 29.26a**). Now we lower the temperature until the superconducting transition occurs. (We assume that the magnitude of $\vec{B}_0$ is not large enough to prevent the phase transition.) What happens to the field?

Measurements of the field outside the sphere show that the field lines become distorted as in Fig. 29.26b. There is no longer any field inside the material, except possibly in a very thin surface layer a hundred or so atoms thick. If a coil is wrapped around the sphere, the emf induced in the coil shows that during the superconducting transition the magnetic flux through the coil decreases from its initial value to zero; this is consistent with the absence of field inside the material. Finally, if the field is now turned off while the material is still in its superconducting phase, no emf is induced in the coil, and measurements show no field outside the sphere (Fig. 29.26c).

We conclude that during a superconducting transition in the presence of the field $\vec{B}_0$, all of the magnetic flux is expelled from the bulk of the sphere, and the magnetic flux Φ_B through the coil becomes zero. This expulsion of magnetic flux is called the *Meissner effect*. As Fig. 29.26b shows, this expulsion crowds the magnetic field lines closer together to the side of the sphere, increasing $\vec{B}$ there.

Superconductor Levitation and Other Applications

The diamagnetic nature of a superconductor has some interesting *mechanical* consequences. A paramagnetic or ferromagnetic material is attracted by a permanent magnet because the magnetic dipoles in the material align with the nonuniform magnetic field of the permanent magnet. (We discussed this in Section 27.7.) For a diamagnetic material

Figure 29.26 A superconducting material (a) above the critical temperature and (b), (c) below the critical temperature.

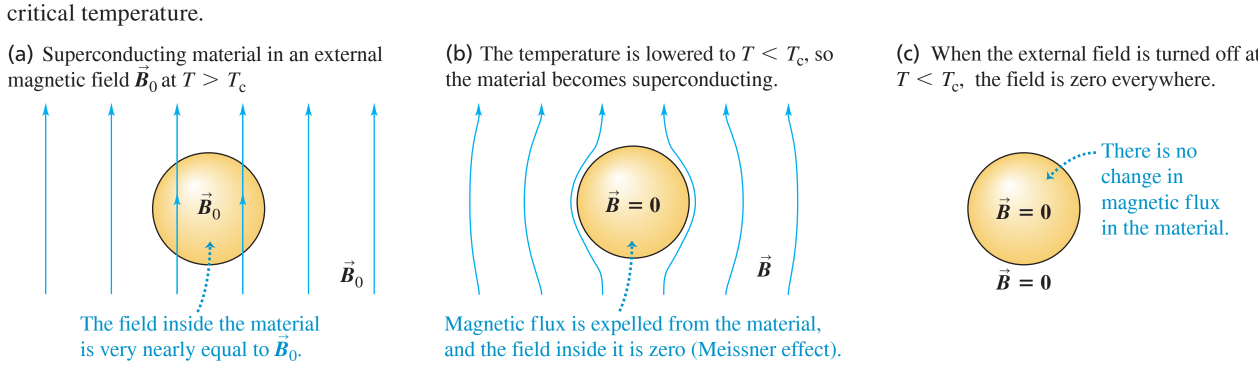
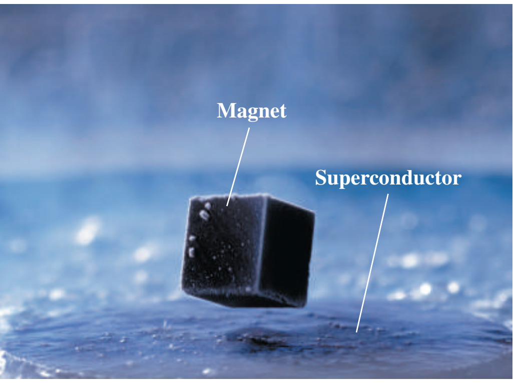


Figure 29.27 A superconductor exerts a repulsive force on a magnet, supporting the magnet in midair.



the magnetization is in the opposite sense, and a diamagnetic material is *repelled* by a permanent magnet. By Newton's third law the magnet is also repelled by the diamagnetic material. **Figure 29.27** shows the repulsion between a specimen of a high-temperature superconductor and a magnet; the magnet is supported ("levitated") by this repulsive magnetic force.

The behavior we have described is characteristic of what are called *type-I superconductors*. There is another class of superconducting materials called *type-II superconductors*. When such a material in the superconducting phase is placed in a magnetic field, the bulk of the material remains superconducting, but thin filaments of material, running parallel to the field, may return to the normal phase. Currents circulate around the boundaries of these filaments, and there *is* magnetic flux inside them. Type-II superconductors are used for electromagnets because they usually have much larger values of B_c than do type-I materials, permitting much larger magnetic fields without destroying the superconducting state. Type-II superconductors have *two* critical magnetic fields: The first, B_{c1}, is the field at which magnetic flux begins to enter the material, forming the filaments just described, and the second, B_{c2}, is the field at which the material becomes normal.

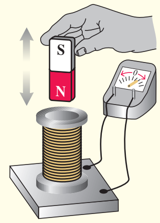
Superconducting electromagnets are in everyday use not only in research laboratories but also in medical MRI (magnetic resonance imaging) scanners. As we described in Section 27.7, scanning a patient through MRI requires a strong magnetic field to align the magnetic dipoles of the patient's atomic nuclei. A steady field of 1.5 T or more is needed, which is very difficult to produce with a conventional electromagnet, since this would require very high currents and hence very high energy losses due to resistance in the electromagnet coils. But with a superconducting electromagnet there is no resistive energy loss, and fields up to 10 T can routinely be attained.

Very sensitive measurements of magnetic fields can be made with superconducting quantum interference devices (SQUIDs), which can detect changes in magnetic flux of less than 10^{-14} Wb; these devices have applications in medicine, geology, and other fields. The number of potential uses for superconductors has increased greatly since the discovery in 1987 of high-temperature superconductors. These materials have critical temperatures that are above the temperature of liquid nitrogen (about 77 K) and so are comparatively easy to attain. Development of practical applications of superconductor science promises to be an exciting chapter in contemporary technology.

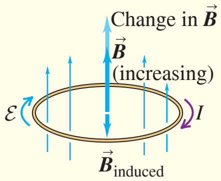
CHAPTER 29 SUMMARY

Faraday's law: Faraday's law states that the induced emf in a closed loop equals the negative of the time rate of change of magnetic flux through the loop. This relationship is valid whether the flux change is caused by a changing magnetic field, motion of the loop, or both. (See Examples 29.1–29.6.)

$$\mathcal{E} = -\frac{d\Phi_B}{dt} \tag{29.3}$$



Lenz's law: Lenz's law states that an induced current or emf always tends to oppose or cancel out the change that caused it. Lenz's law can be derived from Faraday's law and is often easier to use. (See Examples 29.7 and 29.8.)



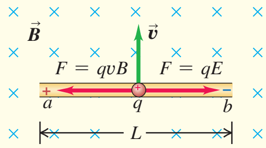
Motional emf: If a conductor moves in a magnetic field, a motional emf is induced. (See Examples 29.9 and 29.10.)

$$\mathcal{E} = vBL \tag{29.6}$$

(conductor with length L moves in uniform $\vec{B}$ field, $\vec{L}$ and $\vec{v}$ both perpendicular to $\vec{B}$ and to each other)

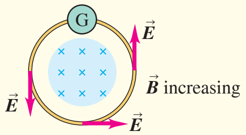
$$\mathcal{E} = \oint (\vec{v} \times \vec{B}) \cdot d\vec{l} \tag{29.7}$$

(all or part of a closed loop moves in a $\vec{B}$ field)



Induced electric fields: When an emf is induced by a changing magnetic flux through a stationary conductor, there is an induced electric field $\vec{E}$ of nonelectrostatic origin. This field is nonconservative and cannot be associated with a potential. (See Example 29.11.)

$$\oint \vec{E} \cdot d\vec{l} = -\frac{d\Phi_B}{dt} \tag{29.10}$$



Displacement current and Maxwell's equations: A time-varying electric field generates displacement current i_D, which acts as a source of magnetic field in exactly the same way as conduction current. The relationships between electric and magnetic fields and their sources can be stated compactly in four equations, called Maxwell's equations. Together they form a complete basis for the relationship of $\vec{E}$ and $\vec{B}$ fields to their sources.

$$i_D = \epsilon \frac{d\Phi_E}{dt} \tag{29.14}$$

(displacement current)

$$\oint \vec{E} \cdot d\vec{A} = \frac{Q_{\text{encl}}}{\epsilon_0} \tag{29.18}$$

(Gauss's law for $\vec{E}$ fields)

$$\oint \vec{B} \cdot d\vec{A} = 0 \tag{29.19}$$

(Gauss's law for $\vec{B}$ fields)

$$\oint \vec{E} \cdot d\vec{l} = -\frac{d\Phi_B}{dt} \tag{29.20}$$

(Faraday's law)

$$\oint \vec{B} \cdot d\vec{l} = \mu_0\left(i_C + \epsilon_0 \frac{d\Phi_E}{dt}\right)_{\text{encl}} \tag{29.21}$$

(Ampere's law including displacement current)

GUIDED PRACTICE

KEY EXAMPLE VARIATION PROBLEMS

Be sure to review EXAMPLES 29.1, 29.2, and 29.3 (Section 29.2) and CONCEPTUAL EXAMPLES 29.7 and 29.8 (Section 29.3) before attempting these problems.

VP29.8.1 A single circular loop of wire with radius 2.40 cm lies in the xy-plane. There is a uniform magnetic field that changes at a steady rate from 0.140 T in the $+z$-direction at $t = 0$ to 0.110 T in the $-z$-direction at $t = 2.00$ s. Take the area vector for the loop to be in the $+z$-direction. Find (a) the magnetic flux through the loop at $t = 0$ and $t = 2.00$ s, and (b) the magnitude of the induced emf in the loop while the field is changing.

VP29.8.2 A circular coil with 455 turns of wire has radius 5.00 cm and resistance 14.5 Ω. The area vector of the coil points in the $+z$-direction. The coil is in a region of uniform magnetic field that points 20.0° from the $+z$-direction. The magnitude of the field is initially 0.0600 T and decreases at a rate of -3.00×10^{-3} T/s. Find the magnitude of (a) the induced emf and (b) the induced current in the coil while the field is changing.

VP29.8.3 A square single loop of wire 4.00 cm on a side with 875 turns lies in the xy-plane. The loop is in a uniform magnetic field that changes at a steady rate from $\vec{B} = (0.200\text{ T})\hat{\imath} + (0.150\text{ T})\hat{k}$ at $t = 0$ to $\vec{B} = (0.300\text{ T})\hat{\imath} + (-0.200\text{ T})\hat{k}$ at $t = 3.00$ s. At the time $t = 2.00$ s, find (a) the magnitude of the induced emf and (b) the direction of the induced current in the coil as seen from a point on the $+z$-axis.

VP29.8.4 A coil with area A and N turns of wire lies in the xy-plane. Its area vector points in the $+z$-direction. It is exposed to a uniform time-varying magnetic field $\vec{B} = (B_0 \sin \omega t)\hat{k}$, where ω is a positive constant. Find (a) the induced emf in the coil and (b) the direction of the induced current in the coil as seen from a point on the $+z$-axis at $t = 0$ and at $t = \pi/\omega$.

Be sure to review EXAMPLE 29.9 (Section 29.4) before attempting these problems.

VP29.9.1 An isolated conducting rod of length 8.00 cm is oriented parallel to the x-axis. It moves in the $+y$-direction at 3.90 m/s in the presence of a uniform magnetic field of magnitude 0.600 T that points in the $-z$-direction. Find (a) the magnitude of the motional emf in the rod and (b) the magnitude and direction of the electric field produced in the rod.

VP29.9.2 In the slidewire generator of Fig. 29.15b, you reverse the rod's motion so that it moves to the left rather than to the right. The moving rod is 5.00 cm long and moves at 2.40 m/s, and the uniform magnetic field has magnitude 0.150 T. If the resistance of the circuit at a given instant is 0.0200 Ω, find the current in the circuit and its direction around the circuit.

VP29.9.3 An isolated conducting rod of length 8.00 cm is oriented parallel to the x-axis. At $t = 0$, end a is at $x = y = z = 0$ and end b is at $x = 8.00$ cm, $y = z = 0$. The rod moves with constant velocity $\vec{v} = (1.50\text{ m/s})\hat{\imath} + (2.00\text{ m/s})\hat{\jmath}$ in a uniform magnetic field $\vec{B} = (-0.120\text{ T})\hat{\imath} + (0.150\text{ T})\hat{k}$. (a) Find the magnitude of the induced emf in the rod. (b) Which end of the rod, a or b, is at higher potential?

VP29.9.4 The magnetic force on the moving rod in a slidewire generator (Fig. 29.15b) has magnitude 0.0600 N at an instant when the resistance of the circuit is 0.750 Ω. The rod is 0.190 m long and the magnetic field (which is perpendicular to the plane of the generator) has magnitude 0.550 T. Find the speed of the rod.

Be sure to review EXAMPLE 29.11 (Section 29.5) before attempting these problems.

VP29.11.1 A long solenoid has a cross-sectional area of 3.00 cm². The current through the windings is decreasing at a rate of 22.0 A/s. A wire loop of radius 3.10 cm is around the solenoid, parallel with its coils, and centered on the axis of the solenoid. The magnitude of the induced emf is 15.0 μV. Find (a) the number of turns per meter in the solenoid and (b) the magnitude of the induced electric field within the loop.

VP29.11.2 A circular wire loop of radius 0.360 cm lies in the xz-plane. There is a uniform magnetic field in the y-direction that decreases at 0.0150 T/s. Find the magnitude of the induced electric field in the wire.

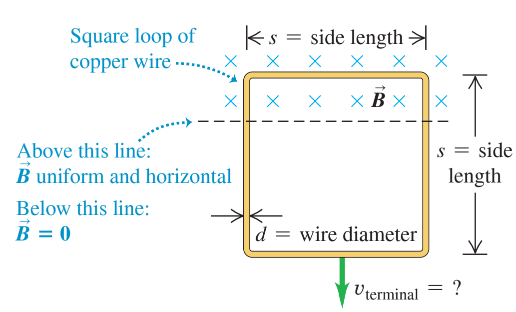
VP29.11.3 A long solenoid with cross-sectional area 4.00 cm² and 965 turns per meter is oriented with its axis along the z-axis. The field inside the solenoid points in the $+z$-direction. A wire loop of radius 5.00 cm is around the solenoid, parallel with its coils, centered on the axis of the solenoid, and lying in the xy-plane. Find the rate of change of the current in the solenoid if the electric field in the loop at the point $x = 5.00$ cm, $y = 0$, $z = 0$ is (a) $\vec{E} = (+1.20 \times 10^{-5}\text{ V/m})\hat{\jmath}$; (b) $\vec{E} = (-1.80 \times 10^{-5}\text{ V/m})\hat{\jmath}$.

VP29.11.4 A long solenoid has 585 turns per meter and a cross-sectional area of 2.70 cm². The current through the windings as a function of time is $(0.600\text{ A/s}^2)t^2$. A wire loop of resistance 0.600 Ω is around the solenoid, parallel with its coils, and centered on the axis of the solenoid. What is the current induced in the loop at $t = 13.9$ s?

BRIDGING PROBLEM A Falling Square Loop

A vertically oriented square loop of copper wire falls from rest in a region in which the field $\vec{B}$ is horizontal, uniform, and perpendicular to the plane of the loop, into a field-free region (**Fig. 29.28**). The side length of the loop is s and the wire diameter is d. The resistivity of copper is ρ_R and the density of copper is ρ_m. If the loop reaches its terminal speed while its upper segment is still in the magnetic-field region, find an expression for the terminal speed.

Figure 29.28 A wire loop falling in a horizontal magnetic field $\vec{B}$. The plane of the loop is perpendicular to $\vec{B}$.



SOLUTION GUIDE

IDENTIFY and SET UP

1. The motion of the loop through the magnetic field induces an emf and a current in the loop. The field then gives rise to a magnetic

force on this current that opposes the downward force of gravity. The loop reaches terminal speed (it no longer accelerates) when the upward magnetic force balances the downward force of gravity.

2. Consider the case in which the entire loop is in the magnetic-field region. Is there an induced emf in this case? If so, what is its direction?
3. Consider the case in which only the upper segment of the loop is in the magnetic-field region. Is there an induced emf in this case? If so, what is its direction?
4. For the case in which there is an induced emf and hence an induced current, what is the direction of the magnetic force on each of the four sides of the loop? What is the direction of the *net* magnetic force on the loop?

EXECUTE

5. For the case in which the loop is falling at speed v and there is an induced emf, find (i) the emf, (ii) the induced current, and (iii) the magnetic force on the loop in terms of its resistance R.
6. Find R and the mass of the loop in terms of the given information about the loop.
7. Use your results from steps 5 and 6 to find an expression for the terminal speed.

EVALUATE

8. How does the terminal speed depend on the magnetic-field magnitude B? Explain why this makes sense.

PROBLEMS

•, ••, •••: Difficulty levels. **CP**: Cumulative problems incorporating material from earlier chapters. **CALC**: Problems requiring calculus. **DATA**: Problems involving real data, scientific evidence, experimental design, and/or statistical reasoning. **BIO**: Biosciences problems.

DISCUSSION QUESTIONS

Q29.1 A sheet of copper is placed between the poles of an electromagnet with the magnetic field perpendicular to the sheet. When the sheet is pulled out, a considerable force is required, and the force required increases with speed. Explain. Is a force required also when the sheet is inserted between the poles? Explain.

Q29.2 In Fig. 29.8, if the angular speed ω of the loop is doubled, then the frequency with which the induced current changes direction doubles, and the maximum emf also doubles. Why? Does the torque required to turn the loop change? Explain.

Q29.3 Two circular loops lie side by side in the same plane. One is connected to a source that supplies an increasing current; the other is a simple closed ring. Is the induced current in the ring in the same direction as the current in the loop connected to the source, or opposite? What if the current in the first loop is decreasing? Explain.

Q29.4 For Eq. (29.6), show that if v is in meters per second, B in teslas, and L in meters, then the units of the right-hand side of the equation are joules per coulomb or volts (the correct SI units for $\mathcal{E}$).

Q29.5 A long, straight conductor passes through the center of a metal ring, perpendicular to its plane. If the current in the conductor increases, is a current induced in the ring? Explain.

Q29.6 A student asserted that if a permanent magnet is dropped down a vertical copper pipe, it eventually reaches a terminal velocity even if there is no air resistance. Why should this be? Or should it?

Q29.7 An airplane is in level flight over Antarctica, where the magnetic field of the earth is mostly directed upward away from the ground. As viewed by a passenger facing toward the front of the plane, is the left or the right wingtip at higher potential? Does your answer depend on the direction the plane is flying?

Q29.8 Consider the situation in Exercise 29.21. In part (a), find the direction of the force that the large circuit exerts on the small one. Explain how this result is consistent with Lenz's law.

Q29.9 A metal rectangle is close to a long, straight, current-carrying wire, with two of its sides parallel to the wire. If the current in the long wire is decreasing, is the rectangle repelled by or attracted to the wire? Explain why this result is consistent with Lenz's law.

Q29.10 A square conducting loop is in a region of uniform, constant magnetic field. Can the loop be rotated about an axis along one side and no emf be induced in the loop? Discuss, in terms of the orientation of the rotation axis relative to the magnetic-field direction.

Q29.11 Example 29.6 discusses the external force that must be applied to the slidewire to move it at constant speed. If there were a break in the left-hand end of the U-shaped conductor, how much force would be needed to move the slidewire at constant speed? As in the example, you can ignore friction.

Q29.12 In the situation shown in Fig. 29.18, would it be appropriate to ask how much *energy* an electron gains during a complete trip around the wire loop with current I'? Would it be appropriate to ask what *potential difference* the electron moves through during such a complete trip? Explain your answers.

Q29.13 A metal ring is oriented with the plane of its area perpendicular to a spatially uniform magnetic field that increases at a steady rate. If the radius of the ring is doubled, by what factor do (a) the emf induced in the ring and (b) the electric field induced in the ring change?

Q29.14 Small one-cylinder gasoline engines sometimes use a device called a *magneto* to supply current to the spark plug. A permanent magnet is attached to the flywheel, and a stationary coil is mounted adjacent to it. Explain how this device is able to generate current. What happens when the magnet passes the coil?

Q29.15 Does Lenz's law say that the induced current in a metal loop always flows to oppose the magnetic flux through that loop? Explain.

Q29.16 Does Faraday's law say that a large magnetic flux induces a large emf in a coil? Explain.

Q29.17 Can one have a displacement current as well as a conduction current within a conductor? Explain.

Q29.18 Your physics study partner asks you to consider a parallel-plate capacitor that has a dielectric completely filling the volume between the plates. He then claims that Eqs. (29.13) and (29.14) show that the conduction current in the dielectric equals the displacement current in the dielectric. Do you agree? Explain.

Q29.19 Match the mathematical statements of Maxwell's equations as given in Section 29.7 to these verbal statements. (a) Closed electric field lines are evidently produced only by changing magnetic flux. (b) Closed magnetic field lines are produced both by the motion of electric charge and by changing electric flux. (c) Electric field lines can start on positive charges and end on negative charges. (d) Evidently there are no magnetic monopoles on which to start and end magnetic field lines.

Q29.20 If magnetic monopoles existed, the right-hand side of Eq. (29.20) would include a term proportional to the current of magnetic monopoles. Suppose a steady monopole current is moving in a long straight wire. Sketch the *electric* field lines that such a current would produce.

Q29.21 A type-II superconductor in an external field between B_{c1} and B_{c2} has regions that contain magnetic flux and have resistance, and also has superconducting regions. What is the resistance of a long, thin cylinder of such material?

EXERCISES

Section 29.2 Faraday's Law

29.1 • A single loop of wire with an area of 0.0900 m^2 is in a uniform magnetic field that has an initial value of 3.80 T, is perpendicular to the plane of the loop, and is decreasing at a constant rate of 0.190 T/s. (a) What emf is induced in this loop? (b) If the loop has a resistance of 0.600 Ω, find the current induced in the loop.

29.2 •• In a physics laboratory experiment, a coil with 200 turns enclosing an area of 12 cm^2 is rotated in 0.040 s from a position where its plane is perpendicular to the earth's magnetic field to a position where its plane is parallel to the field. The earth's magnetic field at the lab location is 6.0×10^{-5} T. (a) What is the magnetic flux through each turn of the coil before it is rotated? After it is rotated? (b) What is the average emf induced in the coil?

29.3 • The magnetic flux through a coil is given by $\Phi_B = \alpha t - \beta t^3$, where α and β are constants. (a) What are the units of α and β? (b) If the induced emf is zero at $t = 0.500$ s, how is α related to β? (c) If the emf at $t = 0$ is -1.60 V, what is the emf at $t = 0.250$ s?

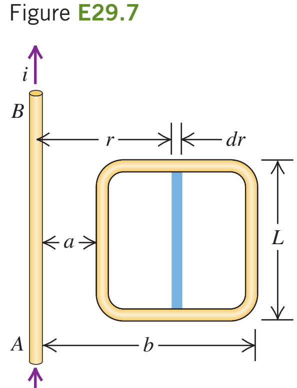
29.4 •• A small, closely wound coil has N turns, area A, and resistance R. The coil is initially in a uniform magnetic field that has magnitude B and a direction perpendicular to the plane of the loop. The coil is then rapidly pulled out of the field so that the flux through the coil is reduced to zero in time Δt. (a) What are the magnitude of the average emf $\mathcal{E}_{av}$ and average current I_{av} induced in the coil? (b) The total charge Q that flows through the coil is given by $Q = I_{av}\Delta t$. Derive an expression for Q in terms of N, A, B, and R. Note that Q does not depend on Δt. (c) What is Q if $N = 150$ turns, $A = 4.50$ cm^2, $R = 30.0\ \Omega$, and $B = 0.200$ T?

29.5 • A circular loop of wire with a radius of 12.0 cm and oriented in the horizontal xy-plane is located in a region of uniform magnetic field. A field of 1.5 T is directed along the positive z-direction, which is upward. (a) If the loop is removed from the field region in a time interval of 2.0 ms, find the average emf that will be induced in the wire loop during the extraction process. (b) If the coil is viewed looking down on it from above, is the induced current in the loop clockwise or counterclockwise?

29.6 • **CALC** A coil 4.00 cm in radius, containing 500 turns, is placed in a uniform magnetic field that varies with time according to $B = (0.0120\ \text{T/s})t + (3.00 \times 10^{-5}\ \text{T/s}^4)t^4$. The coil is connected to a 600 Ω resistor, and its plane is perpendicular to the magnetic field. You can ignore the resistance of the coil. (a) Find the magnitude of the induced emf in the coil as a function of time. (b) What is the current in the resistor at time $t = 5.00$ s?

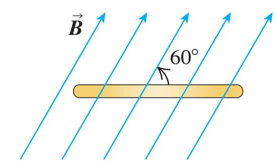
29.7 • **CALC** The current in the long, straight wire AB shown in **Fig. E29.7** is upward and is increasing steadily at a rate di/dt. (a) At an instant when the current is i, what are the magnitude and direction of the field $\vec{B}$ at a distance r to the right of the wire? (b) What is the flux $d\Phi_B$ through the narrow, shaded strip? (c) What is the total flux through the loop? (d) What is the induced emf in the loop? (e) Evaluate the numerical value of the induced emf if $a = 12.0$ cm, $b = 36.0$ cm, $L = 24.0$ cm, and $di/dt = 9.60$ A/s.

Figure **E29.7**



29.8 • **CALC** A flat, circular, steel loop of radius 75 cm is at rest in a uniform magnetic field, as shown in an edge-on view in **Fig. E29.8**. The field is changing with time, according to $B(t) = (1.4\ \text{T})e^{-(0.057\ \text{s}^{-1})t}$. (a) Find the emf induced in the loop as a function of time. (b) When is the induced emf equal to $\frac{1}{10}$ of its initial value? (c) Find the direction of the current induced in the loop, as viewed from above the loop.

Figure **E29.8**



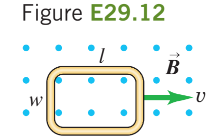
29.9 • **Shrinking Loop.** A circular loop of flexible iron wire has an initial circumference of 165.0 cm, but its circumference is decreasing at a constant rate of 12.0 cm/s due to a tangential pull on the wire. The loop is in a constant, uniform magnetic field oriented perpendicular to the plane of the loop and with magnitude 0.500 T. (a) Find the emf induced in the loop at the instant when 9.0 s have passed. (b) Find the direction of the induced current in the loop as viewed looking along the direction of the magnetic field.

29.10 • A closely wound rectangular coil of 80 turns has dimensions of 25.0 cm by 40.0 cm. The plane of the coil is rotated from a position where it makes an angle of 37.0° with a magnetic field of 1.70 T to a position perpendicular to the field. The rotation takes 0.0600 s. What is the average emf induced in the coil?

29.11 •• **CALC** A circular loop of wire with radius 2.00 cm and resistance 0.600 Ω is in a region of a spatially uniform magnetic field $\vec{B}$ that is perpendicular to the plane of the loop. At $t = 0$ the magnetic field has magnitude $B_0 = 3.00$ T. The magnetic field then decreases according to the equation $B(t) = B_0 e^{-t/\tau}$, where $\tau = 0.500$ s. (a) What is the maximum magnitude of the current I induced in the loop? (b) What is the induced current I when $t = 1.50$ s?

29.12 • A flat, rectangular coil of dimensions l and w is pulled with uniform speed v through a uniform magnetic field B with the plane of its area perpendicular to the field (**Fig. E29.12**). (a) Find the emf induced in this coil. (b) If the speed and magnetic field are both tripled, what is the induced emf?

Figure **E29.12**

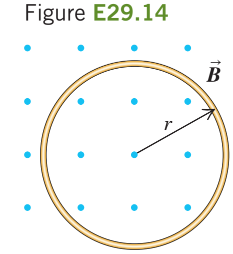


29.13 •• The armature of a small generator consists of a flat, square coil with 120 turns and sides with a length of 1.60 cm. The coil rotates in a magnetic field of 0.0750 T. What is the angular speed of the coil if the maximum emf produced is 24.0 mV?

Section 29.3 Lenz's Law

29.14 • A circular loop of wire with radius $r = 0.0480$ m and resistance $R = 0.160\ \Omega$ is in a region of spatially uniform magnetic field, as shown in **Fig. E29.14**. The magnetic field is directed out of the plane of the figure. The magnetic field has an initial value of 8.00 T and is decreasing at a rate of $dB/dt = -0.680$ T/s. (a) Is the induced current in the loop clockwise or counterclockwise? (b) What is the rate at which electrical energy is being dissipated by the resistance of the loop?

Figure **E29.14**



29.15 • A circular loop of wire is in a region of spatially uniform magnetic field, as shown in **Fig. E29.15**. The magnetic field is directed into the plane of the figure. Determine the direction (clockwise or counterclockwise) of the induced current in the loop when (a) B is increasing; (b) B is decreasing; (c) B is constant with value B_0. Explain your reasoning.

Figure **E29.15**