# Fluid Dynamics

## CHAPTER 18

### 18–1 General Concepts of Fluid Flow

One way of describing the motion of a fluid is to divide the fluid into infinitesimal volume elements, which we may call fluid particles, and to follow the motion of each of these particles. This is a formidable task. We would give coordinates $x$, $y$, $z$ to each such fluid particle and would specify these as functions of the time $t$. The coordinates $x$, $y$, $z$ at the time $t$ of the fluid particle which was at $x_0$, $y_0$, $z_0$ at the time $t_0$ would be determined by functions $x(x_0,y_0,z_0,t_0,t)$, $y(x_0,y_0,z_0,t_0,t)$, $z(x_0,y_0,z_0,t_0,t)$, which then describe the motion of the fluid. This procedure is a direct generalization of the concepts of particle mechanics and was first developed by Joseph Louis Lagrange (1736–1813).

There is a treatment, developed by Leonhard Euler (1707–1783), which is more convenient for most purposes. In it we give up the attempt to specify the history of each fluid particle and instead specify the density and the velocity of the fluid at each point in space at each instant of time. This is the method we shall follow here. We describe the motion of the fluid by specifying the density $\rho(x,y,z,t)$ and the velocity $\mathbf{v}(x,y,z,t)$ at the point $(x,y,z)$ at the time $t$. We thus focus our attention on what is happening at a particular point in space at a particular time, rather than on what is happening to a particular fluid particle. Any quantity used in describing the state of the fluid, for example the pressure $p$, will have a definite value at each point in space and at each instant of time. Although this description of fluid motion focuses attention on a point in space rather than on a fluid particle, we cannot avoid following the fluid particles themselves, at least for short time intervals $dt$. For it is the particles, after all, and not the space points, to which the laws of mechanics apply.

In order to understand the nature of the simplifications we shall make, let us consider first some general characteristics of fluid flow.

Fluid flow can be *steady* or *nonsteady*. When the fluid velocity $\mathbf{v}$ at any given point is constant in time, the fluid motion is said to be steady. That is, at any given point in a steady flow the velocity of each passing fluid particle is always the same. At some other point a particle may travel with a different velocity, but every other particle which passes this second point behaves there just as this particle did when it passed this point. These conditions can be achieved at low flow speeds; a gently flowing stream is an example. In nonsteady flow, as in a tidal bore, the velocities $\mathbf{v}$ *are* a function of the time. In the case of turbulent flow, such as rapids or a waterfall, the velocities vary erratically from point to point as well as from time to time.

Fluid flow can be *rotational* or *irrotational*. If the element of fluid at each point has no net angular velocity about that point, the fluid flow is irrotational. We can imagine a small paddle wheel immersed in the moving fluid (Fig. 18-1). If the wheel moves without rotating, the motion is irrotational; otherwise it is rotational. Irrotational flow is important chiefly because it yields fairly simple mathematical problems. Angular momentum will play no role here and $\mathbf{v}$ is relatively simple. Rotational flow includes vortex motion, such as whirlpools or eddies, and motion in which the velocity vector varies in the transverse direction.

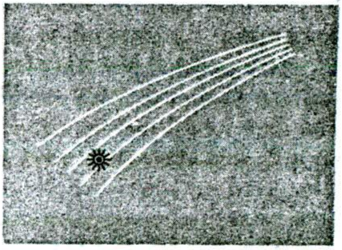


**Fig. 18-1** A small paddle wheel placed in a flowing liquid rotates in rotational flow and does not rotate in irrotational flow.

Fluid flow can be *compressible* or *incompressible*. Liquids can usually be considered as flowing incompressibly. But even a highly compressible gas may sometimes undergo unimportant changes in density. Its flow is then practically incompressible. In flight at speeds much lower than the speed of sound in air (described by subsonic aerodynamics), the motion of the air relative to the wings is one of nearly incompressible flow. In such cases the density $\rho$ is a constant, independent of $x$, $y$, $z$, and $t$, and the mathematical treatment of fluid flow is thereby greatly simplified.

Finally, fluid flow can be *viscous* or *nonviscous*. Viscosity in fluid motion is the analog of friction in the motion of solids. In many cases, such as in lubrication problems, it is extremely important. Sometimes, however, it is negligible. Viscosity introduces tangential forces between layers of fluid in relative motion and results in dissipation of mechanical energy.

We shall confine our discussion of fluid dynamics for the most part to *steady, irrotational, incompressible, nonviscous* flow. The mathematical simplifications resulting should be obvious. We run the danger, however, of making so many simplifying assumptions that we are no longer talking about a meaningfully real fluid.* Furthermore, it is sometimes difficult to decide whether a given property of a fluid—its viscosity, say—can be neglected in a particular situation. In spite of all this the restricted analysis that we are going to give has wide application in practice, as we shall see.

## 18–2 Streamlines

In steady flow the velocity **v** at a given point is constant in time. Consider the point $P$ (Fig. 18–2) within the fluid. Since **v** at $P$ does not change in time, every particle arriving at $P$ will pass on with the same speed in the same direction. The same is true about the points $Q$ and $R$. Hence, if we trace out the path of the particle, as is done in the figure, that curve will be the path of every particle arriving at $P$. This curve is called a *streamline.* A streamline is parallel to the velocity of the fluid particles at every point. No two streamlines can ever cross one another, for if they did, an oncoming fluid particle could go either one way or the other, and the flow could not be steady. In steady flow the pattern of streamlines in a flow is stationary with time.†

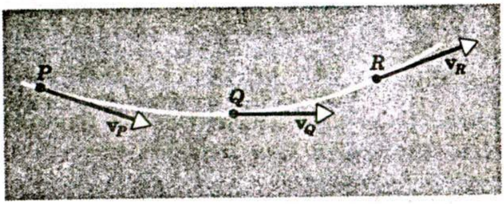


**Fig. 18–2** A particle passing through points $P$, $Q$, and $R$ traces out a streamline, assuming steady flow. Any other particle passing through $P$ must be traveling along the same streamline in steady flow.

In principle we can draw a streamline through every point in the fluid. Let us assume steady flow and select a finite number of streamlines to form a bundle, like the streamline pattern of Fig. 18–3. This tubular region is called a *tube of flow.* The boundary of such a tube consists of streamlines and is always parallel to the velocity of the fluid particles. Hence, *no fluid can cross the boundaries of a tube of flow* and the tube behaves somewhat like a pipe of the same shape. The fluid that enters at one end must leave at the other.

---

* Richard Feynman has pointed out that the late John von Neuman called this idealized fluid "dry water."

† The family of streamlines in a fluid is so drawn that, at any point in the fluid, the direction of the instantaneous velocity **v** for the fluid particle at that point is tangent to the streamline at that point. In nonsteady flow the pattern of streamlines in the fluid changes as time goes on and the path of an individual fluid particle through the fluid does not coincide with a streamline of a given instant. The streamline and the line of motion of the particle touch each other at the point, locating the particle at the instant in question. The path or line of motion and the streamline coincide only for steady flow.

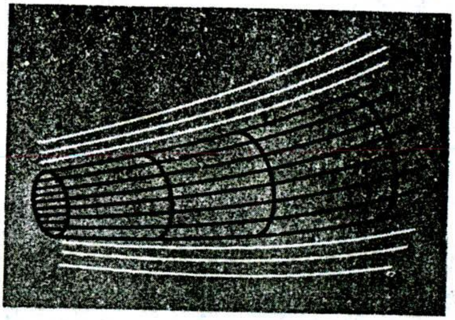


**Fig. 18-3** A tube of flow made up of a bundle of streamlines.

## 18-3 The Equation of Continuity

In Fig. 18-4 we have drawn a thin tube of flow. The velocity of the fluid inside, although parallel to the tube at any point, may have different magnitudes at different points. Let the speed be $v_1$ for fluid particles at $P$ and $v_2$ for fluid particles at $Q$. Let $A_1$ and $A_2$ be the cross-sectional areas of the tubes perpendicular to the streamlines at the points $P$ and $Q$, respectively. In the time interval $\Delta t$ a fluid element travels approximately the distance $v\,\Delta t$. Then the mass of fluid $\Delta m_1$ crossing $A_1$ in the time interval $\Delta t$ is approximately

$$\Delta m_1 = \rho_1 A_1 v_1\,\Delta t$$

or the *mass flux* $\Delta m_1/\Delta t$ is approximately $\rho_1 A_1 v_1$. We must take $\Delta t$ small enough so that in this time interval neither $v$ nor $A$ varies appreciably over the distance the fluid travels. In the limit as $\Delta t \rightarrow 0$, we obtain the precise definitions:

$$\text{mass flux at } P = \rho_1 A_1 v_1, \qquad \text{and}$$

$$\text{mass flux at } Q = \rho_2 A_2 v_2,$$

where $\rho_1$ and $\rho_2$ are the fluid densities at $P$ and $Q$ respectively. Since no fluid can leave through the walls of the tube and there are no "sources" or "sinks" wherein fluid can be created or destroyed in the tube, the mass crossing each section of the tube per unit time must be the same. In par-

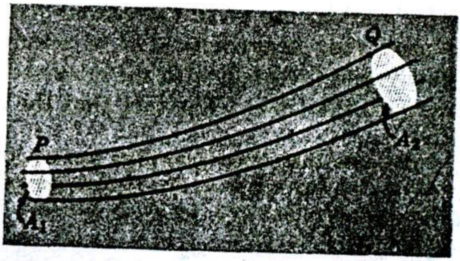


**Fig. 18-4** A tube of flow used in proving the equation of continuity.

ticular, the mass flux at $P$ must equal that at $Q$:

$$\rho_1 A_1 v_1 = \rho_2 A_2 v_2,$$

or

$$\rho A v = \text{constant}. \tag{18–1}$$

This result (Eq. 18–1) expresses the law of conservation of mass in fluid dynamics.

Would you expect Eq. 18–1 to hold when the flow is viscous? When it is rotational?

In the more general case in which sources or sinks are present and in which the density varies with time as well as position, mass must still be conserved and we can write down (without proof) an *equation of continuity* that expresses this fact. It is

$$\frac{\partial(\rho v_x)}{\partial x} + \frac{\partial(\rho v_y)}{\partial y} + \frac{\partial(\rho v_z)}{\partial z} + \frac{\partial \rho}{\partial t} = S \tag{18–2}$$

in which $v_x$, $v_y$, and $v_z$ are the velocity components of the fluid; like the density $\rho$ they vary both with position and with time.*

Let us consider a small volume element in such a fluid. It can be shown that:

1. The sum of the first three terms of Eq. 18–2 gives the net outflow, per unit volume, of mass from the volume element.
2. The fourth term gives the rate, per unit volume, at which mass is accumulating within the volume element.
3. The last term, $S$, gives the rate, per unit volume, at which mass is being introduced into volume element from a "source" (if $S$ is positive) or is disappearing from the volume element into a "sink" (if $S$ is negative).

It is clear that, with these interpretations of its terms, Eq. 18–2 is a statement of the conservation of mass for fluid flow. Is this equation dimensionally correct?

If $S = 0$ in Eq. 18–2, there are no sources or sinks. If the sum of the first three terms is negative there is a net *inflow* of mass to the volume element. Thus the mass contained in the element must increase with time as fluid "piles up." This is in agreement with Eq. 18–2 because, for the conditions stated, $\partial\rho/\partial t$ must be positive, which means that the density of the fluid (and thus the mass of the fluid) in the volume element is increasing as time goes on.

If the fluid is incompressible, as we shall henceforth assume, then $\rho_1 = \rho_2$ and Eq. 18–1 takes on the simpler form

$$A_1 v_1 = A_2 v_2.$$

or

$$A v = \text{constant} \tag{18–3}$$

The product $Av$ gives the *volume flux* or flow rate, as it is often called. Notice that it predicts that in steady incompressible flow the speed of flow varies inversely with the cross-sectional area, being larger in narrower parts of the tube. The fact that the product $Av$ remains constant along

---

* Because these four quantities are functions of more than one variable we have written the derivatives in Eq. 18–2 as partial derivatives. See, for example, *Physical Mechanics*, Section 12.3, by R. B. Lindsay, D. Van Nostrand Company, 1961, for a derivation of Eq. 18–2.

a tube of flow allows us to interpret the streamline picture somewhat. In a narrow part of the tube the streamlines must crowd closer together than in a wide part. Hence, as the distance between streamlines decreases, the fluid speed must increase. Therefore, we conclude that widely spaced streamlines indicate regions of low speed, and closely spaced streamlines indicate regions of high speed.

We can obtain another interesting result by applying the second law of motion to the flow of fluid between $P$ and $Q$ (Fig. 18-4). A fluid particle at $P$ with speed $v_1$ must be decelerated in the forward direction in acquiring the smaller forward speed $v_2$ at $Q$. Hence the fluid is decelerated in going from $P$ to $Q$. The deceleration can come about from a difference in pressure acting on the fluid particle flowing from $P$ to $Q$ or from the action of gravity. In a horizontal tube of flow the gravitational force does not change. Hence we can conclude that in steady horizontal flow the pressure is greatest where the speed is least.

Were you ever in a crowd when it started to push its way through a small opened door? Outside in the back of the crowd the cross-sectional area was large, the pressure was great, but the speed of advance rather small. Through the door of small cross section the pressure was relieved and the speed of advance gratifyingly increased. This particular "human fluid" is compressible and viscous, of course, and the flow is sometimes turbulent and rotational.

## 18-4 Bernoulli's Equation

Bernoulli's equation is a fundamental relation in fluid mechanics. Like all equations in fluid mechanics it is not a new principle but is derivable from the basic laws of Newtonian mechanics. We will find it convenient to derive it from the work-energy theorem (see Section 7-4), for it is essentially a statement of the work-energy theorem for fluid flow.

Consider the nonviscous, steady, incompressible flow of a fluid through the pipeline or tube of flow in Fig. 18-5. The portion of pipe shown in the figure has a uniform cross section $A_1$ at the left. It is horizontal there at an elevation $y_1$ above some reference level. It gradually widens and rises and at the right has a uniform cross section $A_2$. It is horizontal there at an elevation $y_2$. Let us concentrate our attention on the portion of fluid represented by both cross-shading and horizontal shading and call this fluid the "system." Consider then the motion of the system from the position shown in ($a$) to that in ($b$). At all points in the narrow part of the pipe the pressure is $p_1$ and the speed $v_1$; at all points in the wide portion the pressure is $p_2$ and the speed $v_2$.

The work-energy theorem (see Eq. 7-14) states: *The work done by the resultant force acting on a system is equal to the change in kinetic energy of the system.* In Fig. 18-5 the forces that do work on the system, assuming that we can neglect viscous forces, are the pressure forces $p_1A_1$ and $p_2A_2$ that act on the left- and right-hand ends of the system, respectively, and the force of gravity. As fluid flows through the pipe the net effect, as a

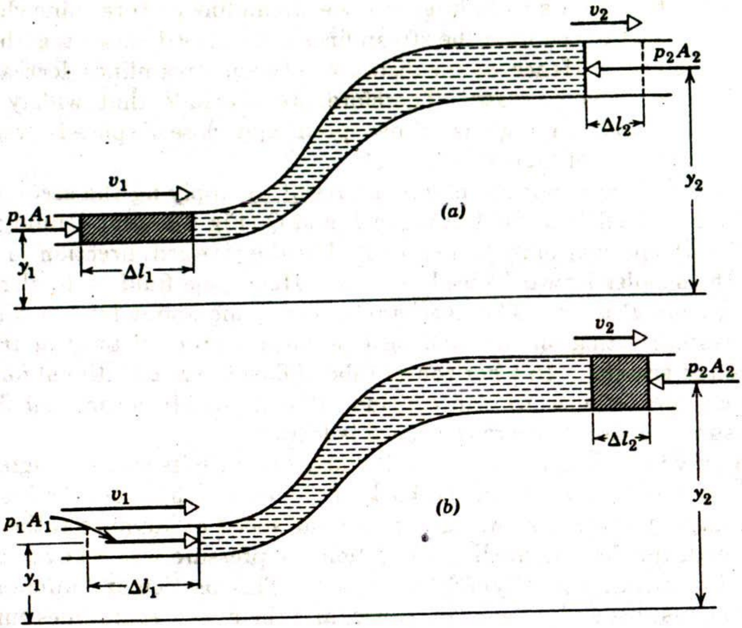


**Fig. 18–5** A portion of fluid (cross-shading *and* horizontal shading) moves through a section of pipeline from the position shown in (*a*) to that shown in (*b*).

comparison of Figs. 18–5*a* and *b* shows, is to raise an amount of fluid represented by the crosshatched area in Fig. 18–5*a* to the position shown in Fig. 18–5*b*. The amount of fluid represented by the horizontal shading is unchanged by the flow.

We can find the work $W$ done on the system by the resultant force as follows:

1. The work done on the system by the pressure force $p_1A_1$ is $p_1A_1\,\Delta l_1$.
2. The work done on the system by the pressure force $p_2A_2$ is $-p_2A_2\,\Delta l_2$. Note that it is negative, which means that positive work is done *by* the system.
3. The work done on the system by gravity is associated with lifting the crosshatched fluid from height $y_1$ to height $y_2$ and is $-mg(y_2 - y_1)$ in which $m$ is the mass of fluid in either crosshatched area. It too is negative because work is done by the system *against* the gravitational force.

The work $W$ done *on* the system by the *resultant* force is found by adding these three terms, or

$$W = p_1A_1\,\Delta l_1 - p_2A_2\,\Delta l_2 - mg(y_2 - y_1).$$

Now $A_1\,\Delta l_1 (= A_2\,\Delta l_2)$ is the volume of the crosshatched fluid element, which we can write as $m/\rho$, in which $\rho$ is the (constant) fluid density.

Recall that the two fluid elements have the same mass, so that in setting $A_1\,\Delta l_1 = A_2\,\Delta l_2$ we have assumed the fluid to be incompressible. With this assumption we have

$$W = (p_1 - p_2)(m/\rho) - mg(y_2 - y_1). \tag{18-4a}$$

The change in kinetic energy of the fluid element is

$$\Delta K = \tfrac{1}{2}mv_2^2 - \tfrac{1}{2}mv_1^2. \tag{18-4b}$$

From the work-energy theorem (Eq. 7-14) we then have

$$W = \Delta K$$

or

$$(p_1 - p_2)(m/\rho) - mg(y_2 - y_1) = \tfrac{1}{2}mv_2^2 - \tfrac{1}{2}mv_1^2, \tag{18-5a}$$

which can be rearranged to read

$$p_1 + \tfrac{1}{2}\rho v_1^2 + \rho g y_1 = p_2 + \tfrac{1}{2}\rho v_2^2 + \rho g y_2. \tag{18-5b}$$

Since the subscripts 1 and 2 refer to *any* two locations along the pipeline, we can drop the subscripts and write

$$p + \tfrac{1}{2}\rho v^2 + \rho g y = \text{constant}. \tag{18-6}$$

Equation 18-6 is called *Bernoulli's equation* for steady, nonviscous, incompressible flow. It was first presented by Daniel Bernoulli (1700-1782) in his *Hydrodynamica* in 1738.

Bernoulli's equation is strictly applicable only to steady flow, the quantities involved being evaluated along a streamline. In our figure the streamline used is along the axis of the pipeline. If the flow is irrotational, however, it can be shown (see Problem 21 for a special case) that the constant in Eq. 18-6 is the same for *all* streamlines.

In a nonviscous incompressible fluid we cannot change the temperature of the fluid by mechanical means. Hence, Bernoulli's equation, as stated above, refers to isothermal (constant temperature) processes. It is possible, however, to change the temperature of a nonviscous *compressible* fluid by mechanical means. We can generalize this equation to include a compressible fluid by adding to the left of Eq. 18-6 a term $u$, which represents the *internal energy* per unit volume of the fluid. This term (and the pressure $p$) will have a value that depends on the temperature.

If the flow is viscous, forces of a frictional nature act on the fluid so that some of the work done that appeared as a change in kinetic energy in the nonviscous case appears now as heat energy in the fluid. We must then write Eq. 18-5a as

$$(p_1 - p_2)(m/\rho) - mg(y_2 - y_1) = \tfrac{1}{2}mv_2^2 - \tfrac{1}{2}mv_1^2 + Q$$

where $Q$ represents the heat energy generated in the viscous flow from point 1 to point 2. In practice, Bernoulli's equation can be modified accordingly by use of empirical corrections for conversion of mechanical energy to heat energy. However, if the pipe is smooth and the diameter is large compared to the length, and if the fluid flows slowly and has a small viscosity, the heat generated is negligible.

Just as the statics of a particle is a special case of particle dynamics, so fluid statics is a special case of fluid dynamics. It should come as no surprise, therefore, that the law of pressure change with height in a fluid at rest is included in Bernoulli's equation as a special case. For let the fluid be at rest; then $v_1 = v_2 = 0$ and Eq. 18–5b becomes

$$p_1 + \rho g y_1 = p_2 + \rho g y_2$$

or

$$p_2 - p_1 = -\rho g(y_2 - y_1),$$

which is the same as Eq. 17–3.

In Eq. 18–6 all terms have the dimension of a pressure (check this). The pressure $p + \rho g h$, which would be present even if there were no flow ($v = 0$), is called the *static pressure;* the term $\frac{1}{2}\rho v^2$ is called the *dynamic pressure.*

## 18–5 Applications of Bernoulli's Equation and the Equation of Continuity

Bernoulli's equation can be used to determine fluid speeds by means of pressure measurements. The principle generally used in such measuring devices is the following: The equation of continuity requires that the speed of the fluid at a constriction increase; Bernoulli's equation then shows that the pressure must fall there. That is, for a horizontal pipe $\frac{1}{2}\rho v^2 + p$ equals a constant; if $v$ increases and the fluid is incompressible, $p$ must decrease. This result was also deduced from dynamic considerations in Section 18–3.

### 1. The Venturi meter

This (Fig. 18–6) is a gauge put in a flow pipe to measure the flow speed of a liquid. A liquid of density $\rho$ flows through a pipe of cross-sectional area $A$. At the throat the area is reduced to $a$ and a manometer tube is attached, as shown. Let the manometer liquid, such as mercury, have a density $\rho'$. By applying Bernoulli's equation and the equation of continuity at points 1 and 2, the student can show that the speed of flow at

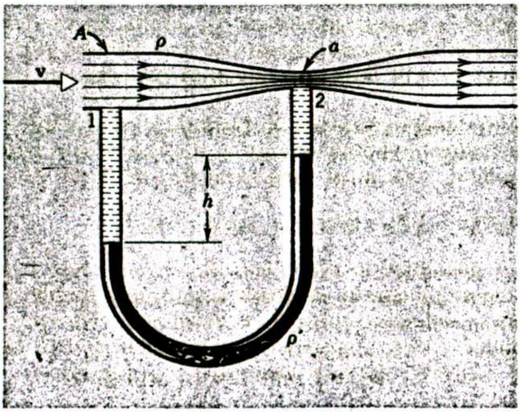


**Fig. 18–6** The Venturi meter, used to measure the speed of flow of a fluid.

$A$ is

$$v = a\sqrt{\frac{2(\rho' - \rho)gh}{\rho(A^2 - a^2)}}.$$

If we want the volume flux or flow rate $R$, which is the volume of liquid transported past any point per second, we simply compute

$$R = vA.$$

### 2. The Pitot tube

This device (Fig. 18–7) is used to measure the flow speed of a gas. Consider the gas, say air, flowing past the openings at $a$. These openings are parallel to the direction of flow and are set far enough back so that the velocity and pressure outside the openings have the free-stream values. The pressure in the left arm of the manometer, which is connected to these openings, is then the static pressure in the gas stream, $p_a$. The opening of the right arm of the manometer is at right angles to the stream. The velocity is reduced to zero at $b$ and the gas is stagnant at that point. The pressure at $b$ is the full ram pressure, $p_b$. Applying Bernoulli's equation to points $a$ and $b$, we obtain

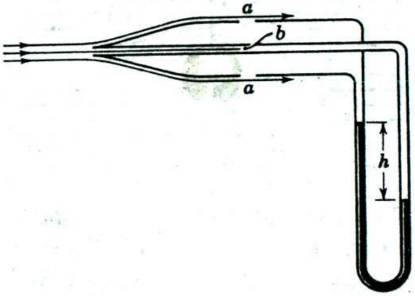


**Fig. 18–7** Cross-sectional diagram of a Pitot tube.

$$p_a + \tfrac{1}{2}\rho v^2 = p_b,$$

where, as shown in the figure, $p_b$ is greater than $p_a$. If $h$ is the difference in height of the liquid in the manometer arms and $\rho'$ is the density of the manometer liquid, then

$$p_a + \rho' gh = p_b.$$

Comparing these two equations, we find

$$\tfrac{1}{2}\rho v^2 = \rho' gh$$

or

$$v = \sqrt{\frac{2gh\rho'}{\rho}},$$

which gives the gas speed. This device can be calibrated to read $v$ directly and is then known as an air-speed indicator.

### 3. Dynamic Lift

Dynamic lift is the force that acts on a body, such as an airplane wing, a hydrofoil, or a spinning baseball, by virtue of its motion through a fluid.

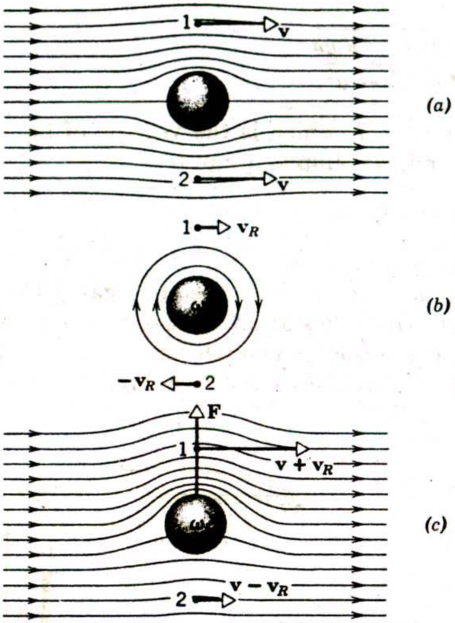


**Fig. 18–8** (*a*) Streamlines for air moving past a (nonrotating) baseball. The velocity **v** is shown for typical corresponding points 1 and 2. (*b*) Streamlines for airflow around a rotating baseball, showing the velocities $\mathbf{v}_R$ and $-\mathbf{v}_R$ at points 1 and 2 respectively. (*c*) the superposition of (*a*) and (*b*). The velocities at points 1 and 2 are shown along with the dynamic lift **F**.

We must distinguish it from *static lift*, which is the buoyant force that acts on an object, such as a balloon, in accord with Archimedes' principle.

Figure 18–8*a* shows the streamlines around a (nonspinning) baseball as it moves through the air. For convenience, we examine the situation from a reference frame in which the baseball is at rest and the air moves past it; this reference frame can be realized in practice by mounting a baseball in a wind tunnel. From the symmetry of the streamlines it is clear that the velocity of the air is the same at corresponding points above and below the ball, such as 1 and 2 in Fig. 18–8*a*. From Bernoulli's equation we then deduce that the pressures at such corresponding points are equal and that the air exerts no upward or downward force on the ball by virtue of its motion; the dynamic lift is zero.

In a separate experiment let us spin a resting baseball about an axis that is perpendicular to the plane of Fig. 18–8*b*. Since the ball is not perfectly smooth, it drags some air around with it, the streamlines associated with this motion being shown in the figure.

Finally, let us combine both motions by throwing a baseball and spinning it at the same time. Figure 18–8*c* shows the resulting streamlines from a reference frame in which the center of mass of the ball is at rest. These streamlines represent a distribution of velocities found by adding (vectorially) at every point the velocities in Figs. 18–8*a* and *b*. The velocities at point 1 add numerically while those at point 2 subtract. Thus the speed at point 1 in Fig. 18–8*c* is *greater* than the speed at point 2 as, indeed, the crowding of the streamlines suggests. From Bernoulli's principle, then, the pressure at point 1 is *less* than the pressure at point 2, so that there is a net upward force (a dynamic lift) on the spinning baseball. Such forces on actual spinning baseballs and tennis balls are well known in practice.*

When dynamic lift on an object occurs it is always associated with an unsymmetrical set of streamlines relatively close together on one side and relatively far apart on the other, like those of Fig. 18–8*c*, that correspond, as Fig. 18–8*b* shows, to circulation of fluid around the object. For the spinning baseball the circulation is obtained by actually spinning the object; in other cases of dynamic lift, of which the airplane wing is a good example, streamline patterns that contain the necessary circulation are obtained by properly shaping the object and properly orienting it in the moving fluid. Figure 18–9 shows streamlines around an airplane wing. As required, they are closer together above the wing than they are below so that (compare Fig. 18–8*c*) Bernoulli's principle predicts the observed upward dynamic lift.

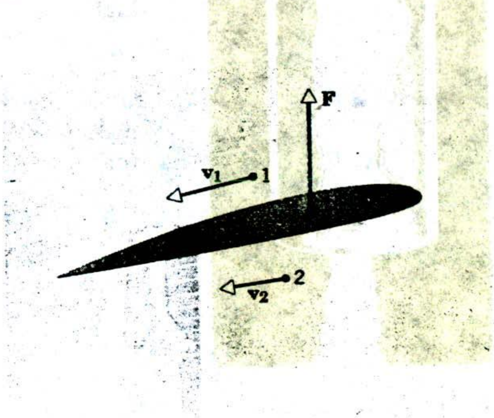


**Fig. 18–9** Streamlines about an airfoil.

### 4. Thrust on a Rocket

As our final example let us compute the thrust on a rocket produced by the escape of its exhaust gases. Consider a chamber (Fig. 18–10) of cross-sectional area $A$ filled with a gas of density $\rho$ at a pressure $p$. Let there be a small orifice of cross-sectional area $A_0$ at the bottom of the chamber. We wish to find the speed $v_0$ with which the gas escapes through the orifice.

Let us write Bernoulli's equation (Eq. 18–5*b*) as

$$p_1 - p_2 = \rho g(y_2 - y_1) + \tfrac{1}{2}\rho(v_2^2 - v_1^2).$$

For a gas the density is so small that we can neglect the variation in pressure with height in a chamber (see Section 17–3). Hence, if $p$ represents the pressure $p_1$ in the chamber and $p_0$ represents the atmospheric pressure

---

* See *The Flettner Ship*, an article by Albert Einstein in his book *Essays in Science*, Philosophical Library, New York.. The Flettner ship, like a sailboat, derives its motive power from the wind. Instead of a sail it has a large cylinder that is caused to rotate about a vertical axis by a small motor. The resulting dynamic "lift" (in this case, horizontal) propels the vessel.

$p_2$ just outside the orifice, we have

$$p - p_0 = \tfrac{1}{2}\rho(v_0^2 - v^2)$$

or

$$v_0^2 = \frac{2(p - p_0)}{\rho} + v^2, \tag{18-7}$$

where $v$ is the speed of the flowing gas inside the chamber and $v_0$ is the *speed of efflux* of the gas through the orifice. Although a gas is compressible and the flow may become turbulent, we can treat the flow as steady and incompressible for pressure and efflux speeds that are not too high.

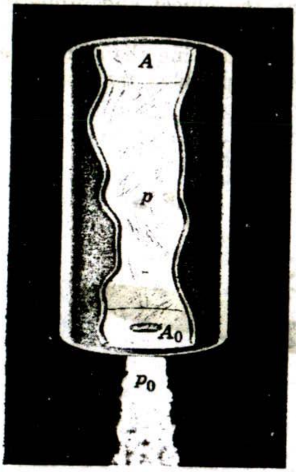


**Fig. 18–10** Fluid streaming out of a chamber.

Now let us assume continuity of mass flow (in a rocket engine this is achieved when the mass of escaping gas equals the mass of gas created by burning the fuel), so that (for an assumed constant density)

$$Av = A_0 v_0.$$

If the orifice is very small so that $A_0 \ll A$, then $v_0 \gg v$, and we can neglect $v^2$ compared to $v_0^2$ in Eq. 18–7. Hence, the speed of efflux is

$$v_0 = \sqrt{\frac{2(p - p_0)}{\rho}}. \tag{18-8}$$

If our chamber is the exhaust chamber of a rocket, the thrust on the rocket (Section 9–7) is $v_0\, dM/dt$. But the mass of gas flowing out in time $dt$ is $dM = \rho A_0 v_0\, dt$, so that

$$v_0 \frac{dM}{dt} = v_0 \rho A_0 v_0 = \rho A_0 v_0^2,$$

and from Eq. 18–8 the thrust is

$$2A_0(p - p_0). \tag{18-9}$$

## 18–6 Conservation of Momentum in Fluid Mechanics

In Newtonian particle mechanics the derivation of the laws of conservation of linear momentum and angular momentum makes explicit use of Newton's third law of motion. The internal forces and torques in a mechanical system cancel one another because of this third law, leaving only the external forces and torques to contribute to the momenta. In the case of a fluid the internal forces are represented by the pressure within the fluid. But the very concept of pressure itself contains Newton's third law implicitly. The force produced by pressure exerted in one direction across any surface element is equal and opposite to the force exerted in the opposite direction across the same surface element. Also, each of these two forces is applied at the same place, namely at the surface element. Both forces must have the same line of action. Hence, in the equations for the time rate of change of linear momentum or of angular momentum of a fluid, the internal pressures will cancel out. We can conclude then that the time rate of change of the total linear momentum in a volume $V$ of moving fluid is equal to the total *external force* acting on it. Likewise, the time rate of change of the total angular momentum in a volume $V$ of moving fluid is equal to the total *external torque* acting on it. The conservation laws of linear and angular momentum follow.

## 18-7 Fields of Flow

In the chapter on gravitation we saw how to summarize the physical state of affairs near masses by use of a field. Each point in the field can be regarded as having a vector associated with it, namely **g**, the gravitational force per unit mass at that point. Or, alternately, we can associate a scalar quantity with each point in space, namely the gravitational potential $V$. We can then draw a surface, called an equipotential surface, through all points that have the same potential. We draw several such surfaces, the potential on one differing by a constant amount from that on the next one, etc. The gravitational force at any point is then directed along a line passing through this point perpendicular to these surfaces, and its magnitude is determined from the rate of change of potential with distance in this direction, as indicated by the spacing and orientation of the equipotential surfaces. By drawing in lines of force we can picture vividly how space is affected by the presence of mass.

Likewise, in fluid dynamics we can summarize the physical state of affairs within a moving fluid by means of a field of flow. In general, the field of flow is a *vector* field. We associate a vector quantity with each point in space, namely the flow velocity **v** at that point. For a steady flow the field of flow is stationary. Of course, even in this case a particular fluid particle may still have a variable velocity as it moves from point to point in the field. The field gives the properties of the space from which we deduce the behavior of particles in that space. If the flow is irrotational, as well as steady, we call it *potential flow*. Then the flow velocity **v** can be related to a velocity potential $\psi$, just as in gravitation **g** can be related to the gravitational potential $V$. If we draw in surfaces of equal velocity potential, as we drew in surfaces of equal gravitational potential, we can deduce **v** from the equipotential flow surfaces just as **g** is deduced from the equipotential gravitational surfaces. Hence, a field for potential flow is analogous to a conservative force field.

A flowing fluid mass can always be divided into tubes of flow. When the flow is steady, the tubes remain unchanged in shape and the fluid that is at one instant in a tube remains inside this tube thereafter. We have seen that the flow velocity inside a tube of flow is parallel to the tube and has a magnitude inversely proportional to the area of the cross section (Eq. 18-1). Let us assign such cross sections to the tubes that the constant of proportionality is the same for all of them; if possible we take this constant to be unity. That is, the volume flux is the same for all tubes, namely unit flux. Then the magnitude of the flow velocity can be

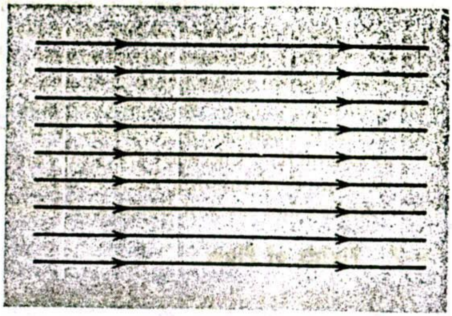


**Fig. 18–11** Streamlines (horizontal) and surfaces of equal velocity potential (vertical) for a homogeneous field of flow.

determined from the areas of the cross sections of the tubes of flow. There is another procedure equivalent to this which consists of setting up a unit area perpendicular to the direction of flow and drawing through it just as many streamlines as the number of units of magnitude of the velocity at that point.

Let us consider some examples of fields of flow. For drawing purposes we consider only *two-dimensional* examples. In these the flow velocity is the same at all points on a line perpendicular to the plane at any point.

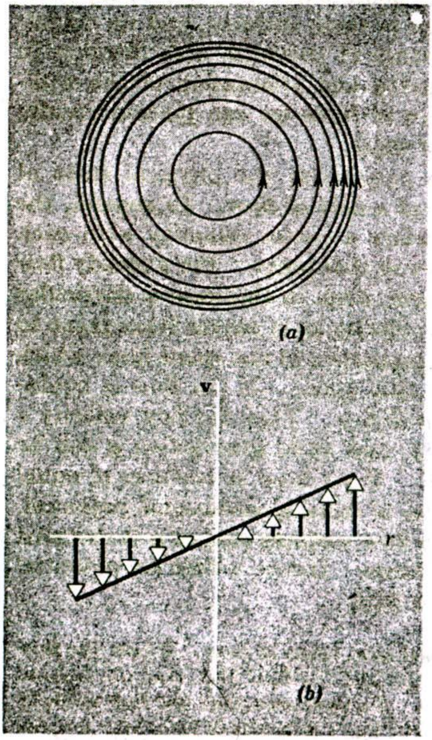


**Fig. 18–12** (*a*) Uniform rotational field of flow. (*b*) Variation of fluid velocity from the center.

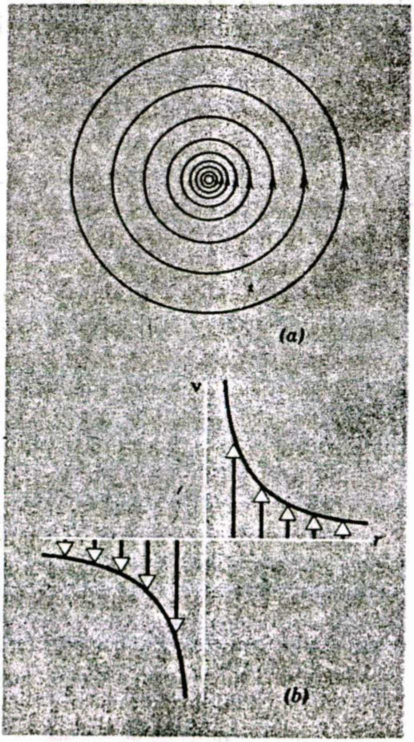


**Fig. 18–13** (*a*) Vortical field of flow. (*b*) Variation of fluid velocity from the center.

In Fig. 18–11 we have drawn a *homogeneous field of flow*. Here all the streamlines are parallel and the flow velocity **v** is the same at all points. We have seen that there are two equivalent ways of deriving the relative magnitudes of the flow velocities from such fields of flow: (*a*) from the widths of the tubes of flow and (*b*) from the distances between lines of equal velocity potential. The latter method applies to steady irrotational flow only. For such flows we draw in the lines of equal velocity potential as dashed lines.

In Fig. 18–12 we show the field for a *uniform rotation* (see Problem 22, Chapter 17). Here $v$ is proportional to $r$. In Fig. 18–13 we draw the field of flow of a *vortex*. In this case $v$ is proportional to $1/r$ (see Problem 19). Notice that both uniform rotation and vortex motion are represented by circular streamlines but are entirely different kinds of flow. Obviously, the *shapes* of the streamlines give only limited information; their spacing is needed too.

Figure 18–14 represents the field of flow for a *source*. All streamlines are directed radially outward. The source is a line through the center perpendicular to the paper emitting a mass per unit time $Q$. The field of flow around a linear *sink* is the same as the source except for the sign of the flow, which is directed radially inward.

For a linear source and linear sink which have the same strengths, $Q$ and $-Q$, and are slightly separated, we obtain the combined field called *linear dipole flow*, shown in Fig. 18–15.

As we shall see later the electrostatic field, the magnetic field, and the field of flow for an electric current are also vector fields. In this connection, the homo-

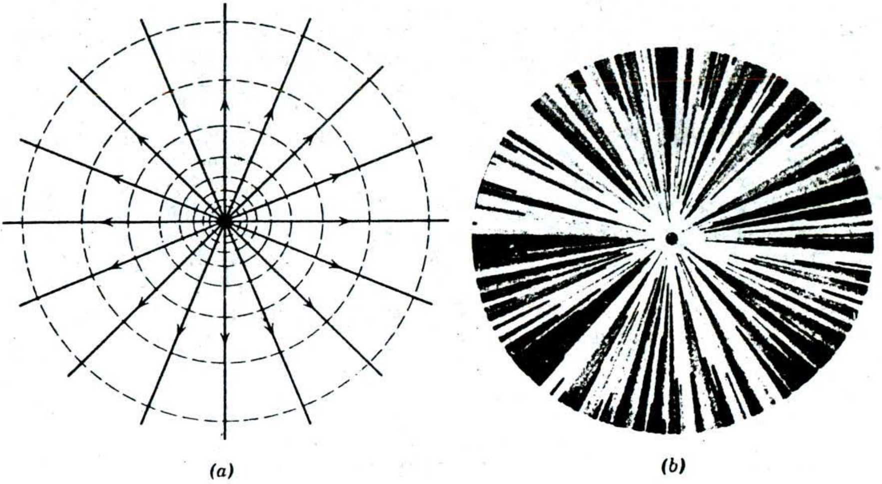


**Fig. 18–14** (*a*) Flow from a linear point source. (*b*) Fluid flow map of the same. The map in this figure is made by allowing water to flow between a horizontal layer of plate glass and a horizontal layer of plaster. In (*b*) the water comes up through a hole in the center of the plaster and flows out toward the edges. The direction of the flow is made visible by sprinkling the plaster with potassium permanganate crystals which dissolve and color the water a deep red. (The fluid flow map was made and photographed by Professor A. D. Moore at the University of Michigan, and is taken from *Introduction to Electric Fields*, by W. E. Rogers, McGraw-Hill Book Co., 1954.)

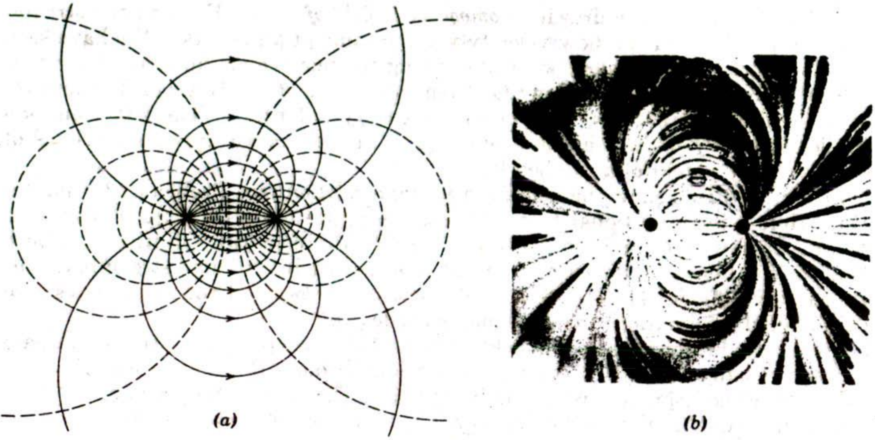


**Fig. 18–15** (*a*) Linear dipole flow. The source is on the left, the sink on the right. (*b*) A fluid flow map of the same. (The fluid flow map was made and photographed by Professor A. D. Moore at the University of Michigan, and is taken from *Introduction to Electric Fields*, by W. E. Rogers, McGraw-Hill Book Co., 1954.)

geneous field (Fig. 18–11) corresponds to the electric field of a plane capacitor, the source field or sink field (Fig. 18–14) correspond to the electric field of a cylindrical capacitor or straight wire of positive or negative charge respectively, and the linear dipole field (Fig. 18–15) corresponds to the electric field of two oppositely charged wires. In all these the field of flow is potential flow and the electric fields are conservative.

The homogeneous field of Fig. 18–11 also represents the magnetic field inside a solenoid. The vortex field of Fig. 18–13 represents the magnetic field around a straight current-carrying wire. This last is an example of a field that is rotational (about the vortex axis).

Because of these analogies between fluid and electromagnetic fields, we can often determine a field of flow, which is impossible to calculate by present mathematical methods, by experimental measurements on appropriate electrical devices.

As we have seen throughout this chapter, the basic field ideas and conservation principles find application in many areas of physics. We shall encounter them many times again.

## QUESTIONS

**1.** Can you assign a coefficient of static friction between two surfaces, one of which is a fluid surface?

**2.** Describe the forces acting on an element of fluid as it flows through a pipe of nonuniform cross section.

**3.** The height of the liquid in the standpipes indicates that the pressure drops along the channel, even though the channel has a uniform cross section and the flowing liquid is incompressible (Fig. 18–16). Explain.

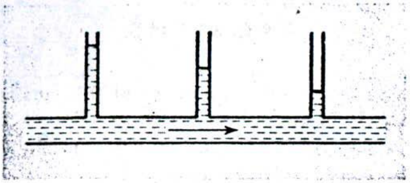


Fig. 18–16

**4.** It is found that liquid will flow faster and more smoothly from a sealed can when two holes are punctured in the can than when one hole is made. Explain.

**5.** In a lecture demonstration a ping-pong ball is kept in midair by a vertical jet of air. Is the equilibrium stable, unstable, or neutral? Explain.

**6.** (*a*) Explain how a pitcher can make a baseball curve to his right or left. Justify your answer by drawing a diagram of the streamlines and applying Bernoulli's equation. (*b*) Why is it easier to throw a curve with a tennis ball than with a baseball?

**7.** Two rowboats moving parallel to one another in the same direction are pulled toward one another. Two automobiles moving parallel are also pulled together. Explain such phenomena on the basis of Bernoulli's equation.

**8.** Can the action of a parachute in retarding free fall be explained by Bernoulli's equation?

**9.** Liquid is flowing inside a horizontal pipe which has a constriction along its length. Vertical tube manometers are attached at both the wide portion and the narrow portion of the pipe. If a stopcock at the exit end is closed, will the liquid in the manometer tubes rise or fall? Explain.

**10.** Can you explain why water flows in a continuous stream down a vertical pipe, whereas it breaks into drops when falling freely?

**11.** Can you explain why an object falling from a great height reaches a steady terminal speed?

**12.** On take off would it be better for an airplane to move into the wind or with the wind? On landing . . . ?

**13.** Does the difference in pressure between the lower and upper surfaces of an airplane wing depend on the altitude of the moving plane? Explain.

**14.** The accumulation of ice on an airplane wing may change its shape in such a way that its lift is greatly reduced. Explain.

**15.** How is an airplane able to fly upside down?

**16.** Why does the factor "2" appear in Eq. 18–9, rather than "1"? One might naively expect that the thrust would simply be the pressure difference times the area, that is, $A_0(p - p_0)$.

**17.** The destructive effect of a tornado (twister) is greater near the center of the disturbance than near the edge. Explain.

**18.** When a stopper is pulled from a filled basin, the water drains out while circulating like a small whirlpool. The angular velocity of a fluid element about a vertical axis through the orifice appears to be greatest near the orifice. Explain.

**19.** Use the criterion of the paddle wheel (Fig. 18–1) to determine which flow fields (Figs. 18–11 through 18–15) are rotational.

**20.** In steady flow the velocity vector $\mathbf{v}$ at any point is constant. Can there then be accelerated motion of the fluid particles? Discuss.

**21.** How can we justify applying Bernoulli's equation to the spinning baseball of Fig. 18-8*c*? Points 1 and 2 are not on the same streamline.

## PROBLEMS

**1.** A garden hose having an internal diameter of 0.75 in. is connected to a lawn sprinkler that consists merely of an enclosure with 24 holes, each 0.050 in. in diameter. If the water in the hose has a speed of 3.0 ft/sec, at what speed does it leave the sprinkler holes?

**2.** Models of torpedoes are sometimes tested in a pipe of flowing water, much as a wind tunnel is used to test model airplanes. Consider a circular pipe of internal diameter 10 in. and a torpedo model, aligned along the axis of the pipe, with a diameter of 2.0 in. The torpedo is to be tested with water flowing past it at 8.0 ft/sec. (*a*) With what speed must the water flow in the unconstricted part of the pipe? (*b*) What will the pressure difference be between the constricted and unconstricted parts of the pipe?

**3.** How much work is done by pressure in forcing 50 $\text{ft}^3$ of water through a 0.50-in. pipe if the difference in pressure at the two ends of the pipe is 15 $\text{lb/in.}^2$?

**4.** Water falls from a height of 60 ft at the rate of 500 $\text{ft}^3$/min and drives a water turbine. What is the maximum power that can be developed by this turbine?

**5.** By applying Bernoulli's equation and the equation of continuity to points 1 and 2 of Fig. 18–6, show that the speed of flow at the entrance is

$$v = a\sqrt{\frac{2(\rho' - \rho)gh}{\rho(A^2 - a^2)}}.$$

**6.** A Venturi meter has a pipe diameter of 10 in. and a throat diameter of 5.0 in. If the water pressure in the pipe is 8.0 $\text{lb/in.}^2$ and in the throat is 6.0 $\text{lb/in.}^2$, determine the rate of flow of water in $\text{ft}^3$/sec (volume flux).

**7.** Consider the Venturi tube of Fig. 18–6 without the manometer. Let $A$ equal $5a$. Suppose the pressure at $A$ is 2.0 atm. Compute the values of $v$ at $A$ and $v'$ at $a$ that would make the pressure $p'$ at $a$ equal to zero. Compute the corresponding volume flow rate if the diameter at $A$ is 5.0 cm. The phenomenon at $a$ when $p'$ falls to nearly zero is known as *cavitation*. The water vaporizes into small bubbles.

**8.** In a horizontal oil pipeline of constant cross-sectional area the pressure decrease between two points 1000 ft apart is 5 $\text{lb/in.}^2$. What is the energy loss per cubic foot of oil per unit distance?

**9.** Figure 18–17 shows liquid discharging from an orifice in a large tank at a distance $h$ below the water level. (*a*) Apply Bernoulli's equation to a streamline connecting points 1, 2, and 3, and show that the speed of efflux is

$$v = \sqrt{2gh}.$$

This is known as Torricelli's law. (*b*) If the orifice were curved directly upward, how high would the liquid stream rise? (*c*) How would viscosity or turbulence affect the analysis?

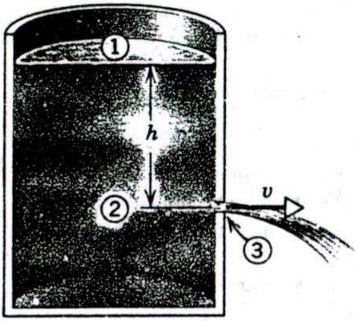


Fig. 18–17

**10.** Suppose that two tanks, each with a large opening at the top, contain different liquids. A small hole is made in the side of each tank at the same depth $h$ below the liquid surface, but one hole has twice the cross-sectional area of the other. (*a*) What is the ratio of the densities of the fluids if it is observed that the mass flux is the same for each hole? (*b*) How do the flow rates (volume flux) from each hole compare? (*c*) Could the flow rates be made equal? How?

**11.** A tank is filled with water to a height $H$. A hole is punched in one of the walls at a depth $h$ below the water surface (Fig. 18–18). (*a*) Find the distance $x$ from the foot of the wall at which the stream strikes the floor. (*b*) Could a hole be punched at another depth so that this second stream would have the same range? If so, at what depth?

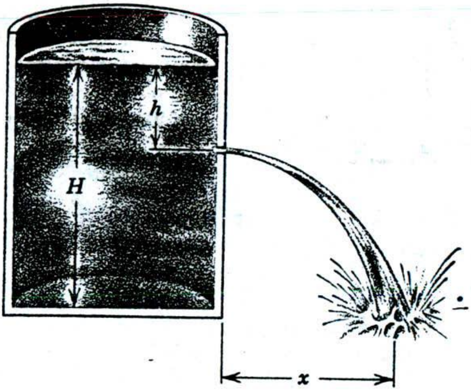


Fig. 18–18

**12.** The upper surface of water in a standpipe is a height $H$ above level ground. At what depth $h$ should a small hole be put to make the emerging horizontal water stream strike the ground at the maximum distance from the base of the standpipe? What is this maximum distance?

**13.** Calculate the speed of efflux of a liquid from an opening in a tank, taking into account the velocity of the top surface of the liquid, as follows. (*a*) Show, from Bernoulli's equation, that

$$v_0^2 = \frac{2gh}{1 - v^2/v_0^2}$$

where $v$ is the speed of the top surface. (*b*) Then consider the flow as one big tube of flow and obtain $v/v_0$ from the equation of continuity, so that

$$v_0 = \sqrt{2gh/[1 - (A_0/A)^2]}$$

where $A$ is the tube cross section at the top and $A_0$ is the tube cross section at the opening. (*c*) Then show that if the hole is small compared to the area of the surface,

$$v_0 \cong \sqrt{2gh}\,[1 + \tfrac{1}{2}(A_0/A)^2].$$

**14.** A Pitot tube is mounted on an airplane wing to determine the speed of the plane relative to the air. The tube contains alcohol and indicates a level difference of 4.9 in. What is the plane's speed in miles/hr relative to the air?

**15.** Air streams horizontally past an airplane wing of area 36 ft$^2$ weighing 540 lb. The speed over the top surface is 200 ft/sec and 150 ft/sec under the bottom surface. What is the lift on the wing? The net force on it?

**16.** If the speed of flow past the lower surface of a wing is 350 ft/sec, what speed of flow over the upper surface will give a lift of 20 lb/ft$^2$?

**17.** (*a*) Consider the stagnant air at the front edge of a wing and the air rushing over the wing surface at a speed $v$. Find the greatest value possible for $v$ in streamline flow, assuming air is incompressible and using Bernoulli's equation. Take the density of air to be $1.2 \times 10^{-3}$ gm/cm$^3$. (*b*) How does this compare with the speed of sound of 770 miles/hr? Can you explain the difference? Why should there be any connection between these quantities?

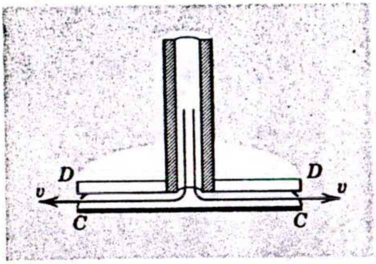


Fig. 18–19

**18.** A hollow tube has a disk $DD$ attached to its end. When air is blown through the tube, the disk attracts the card $CC$. Let the area of the card be $A$ and let $v$ be the average airspeed between $CC$ and $DD$ (Fig. 18–19); calculate the resultant upward force on $CC$. Neglect the card's weight.

**19.** Before Newton proposed his theory of gravitation, a model of planetary motion proposed by René Descartes was widely accepted. In Descartes' model the planets were caught in and dragged along by a whirlpool of ether particles centered around the sun. Newton showed that this vortex scheme contradicted observations, for: (*a*) The speed of an ether particle in the vortex varies inversely as its distance from the sun. (*b*) The period of revolution of such a particle varies directly as the square of its distance from the sun. (*c*) This result contradicts Kepler's third law. Prove (*a*), (*b*), and (*c*).

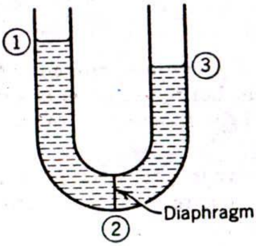


Fig. 18–20

**20.** Consider a uniform U-tube with a diaphragm at the bottom and filled with a liquid to different heights in each arm (see Fig. 18–20). Now imagine that the diaphragm is punctured so that the liquid flows from left to right. (*a*) Show that application of Bernoulli's principle to points 1 and 3 leads to a contradiction. (*b*) Explain why Bernoulli's principle is not applicable here. (*Hint:* Is the flow steady?)

**21.** Show that the constant in Bernoulli's equation (Eq. 18–6) is the same for *all* streamlines in the case of the steady, irrotational flow of Fig. 18–11.

**22.** (*a*) Consider a stream of fluid of density $\rho$ with speed $v_1$ passing *abruptly* from a cylindrical pipe of cross-sectional area $a_1$ into a wider cylindrical pipe of cross-sectional area $a_2$ (see Fig. 18–21). The jet will mix with the surrounding fluid and, after the mixing, will flow on almost uniformly with an average speed $v_2$. Without referring to

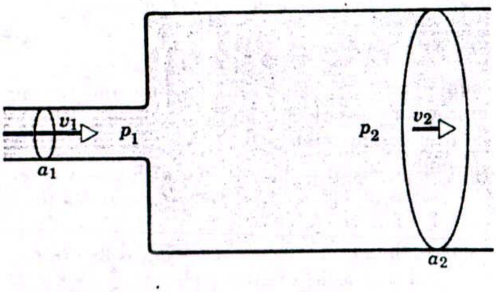


Fig. 18–21

the details of the mixing, use momentum ideas to show that the increase in pressure due to the mixing is approximately

$$p_2 - p_1 = \rho v_2(v_1 - v_2).$$

(*b*) Show from Bernoulli's principle that in a *gradually* widening pipe we would get

$$p_2 - p_1 = \tfrac{1}{2}\rho(v_1^2 - v_2^2)$$

and explain the loss of pressure [the difference is $\frac{1}{2}\rho(v_1 - v_2)^2$] due to the abrupt enlargement of the pipe. Can you draw an analogy with elastic and inelastic collisions in particle mechanics?

**23.** A force field is conservative if $\oint \mathbf{F} \cdot d\mathbf{s} = 0$. The circle on the integration sign means that the integration is to be taken along a closed curve (a round trip) in the field. A flow is a potential flow (hence irrotational) if $\oint \mathbf{v} \cdot d\mathbf{s} = 0$ for every closed path in the field.

Using this criterion, show that the fields of Figs. 18–11 and 18–14 are fields of potential flow.

**24.** The so-called Poiseuille field of flow is shown in Fig. 18–22. The spacing of the streamlines indicates that although the motion is rectilinear, there is a velocity gradient in the transverse direction. Show that such a flow is rotational.

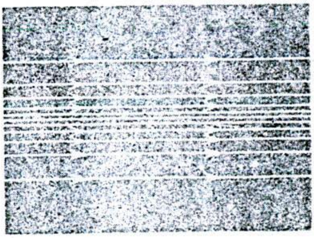


Fig. 18–22

**25.** In flows that are sharply curved centrifugal effects are appreciable. Consider an element of fluid which is moving with speed $v$ along a streamline of a curved flow in a horizontal plane (Fig. 18–23).

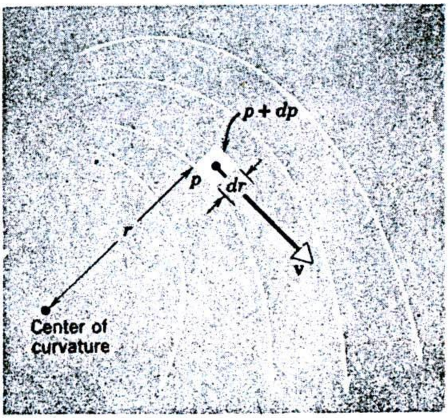


Fig. 18–23

(*a*) Show that $dp/dr = \rho v^2/r$, so that the pressure increases by an amount $\rho v^2/r$ per unit distance perpendicular to the streamline as we go from the concave to the convex side of the streamline.

(*b*) Then use Bernoulli's equation and this result to show that $vr$ equals a constant, so that speeds increase toward the center of curvature. Hence, streamlines that are uniformly spaced in a straight pipe will be crowded toward the inner wall of a curved passage and widely spaced toward the outer wall. This problem should be compared to Problem 17–22 in which the curved motion is produced by rotating a container. There the speed varied directly with $r$, but here it varies inversely.

(*c*) Show that this flow is irrotational.

# Waves in Elastic Media

## CHAPTER 19

### 19–1 Mechanical Waves

Wave motion appears in almost every branch of physics. We are all familiar with water waves. There are also sound waves, as well as light waves, radio waves, and other electromagnetic waves. One formulation of the mechanics of atoms and subatomic particles is called wave mechanics. Clearly the properties and behavior of waves are very important in physics.

In this chapter and the next we confine our attention to waves in deformable or elastic media. These waves, among which ordinary sound waves in air are one example, might be called *mechanical waves.* They originate in the displacement of some portion of an elastic medium from its normal position, causing it to oscillate about an equilibrium position. Because of the elastic properties of the medium, the disturbance is transmitted from one layer to the next. This disturbance, or wave, consequently progresses through the medium. Note that the medium itself does not move as a whole along with the wave motion; the various parts of the medium oscillate only in limited paths. For example, in water waves small floating objects like corks show that the actual motion of various parts of the water is slightly up and down and back and forth. Yet the water waves move steadily along the water. As they reach floating objects they set them in motion, thus transferring energy to them.* Energy can be transmitted over considerable distances by wave motion. The energy in the waves is the kinetic and potential energy of the matter, but the transmission of the

---

* See "Ocean Waves," by Willard Bascom, *Scientific American,* August 1957.

energy comes about by its being passed along from one part of the matter to the next, not by any long-range motion of the matter itself. Mechanical waves are characterized by the transport of energy through matter by the motion of a disturbance in that matter without any corresponding bulk motion of the matter itself.

It is necessary to have a material medium to transmit mechanical waves. We do not need such a medium, however, to transmit electromagnetic waves, light passing freely, for example, through the near vacuum of space from the stars. The properties of the medium that determine the speed of a wave through that medium, as we will see in Section 19–5, are its inertia and its elasticity. All material media, including, say, air, water and steel, possess these properties and can transmit mechanical waves. It is the elasticity that gives rise to the restoring forces on any part of the medium displaced from its equilibrium position; it is the inertia that tells us how this displaced portion of the medium will respond to these restoring forces. Together these two factors determine the wave speed.

### 19–2 Types of Waves

In listing water waves, light waves, and sound waves as examples of wave motion, we are classifying waves according to their broad physical properties. Waves can be classified in other ways.

We can distinguish different kinds of waves by considering how the motions of the particles of matter are related to the direction of propagation of the waves themselves. If the motions of the matter particles conveying the wave are perpendicular to the direction of propagation of the wave itself, we then have a *transverse* wave. For example, when a vertical string under tension is set oscillating back and forth at one end, a transverse wave travels down the string; the disturbance moves along the string but the string particles vibrate at right angles to the direction of propagation of the disturbance (Fig. 19–1*a*).

Light waves are not mechanical waves. The disturbance that travels along is not a motion of matter but an electromagnetic field (Chapter 40). But because the electric and magnetic fields are perpendicular to the direction of propagation, light waves are also transverse waves.

If, however, the motion of the particles conveying a mechanical wave is back and forth along the direction of propagation, we then have a *longitudinal wave*. For example, when a vertical spring under tension is set oscillating up and down at one end, a longitudinal wave travels along the spring; the coils vibrate back and forth in the direction in which the disturbance travels along the spring (Fig. 19–1*b*). Sound waves in a gas are longitudinal waves. We shall discuss them in greater detail in Chapter 20.

Some waves are neither purely longitudinal nor purely transverse. For example, in waves on the surface of water the particles of water move both up and down and back and forth, tracing out elliptical paths as the water waves move by.

Waves can also be classified as one-, two-, and three-dimensional waves,

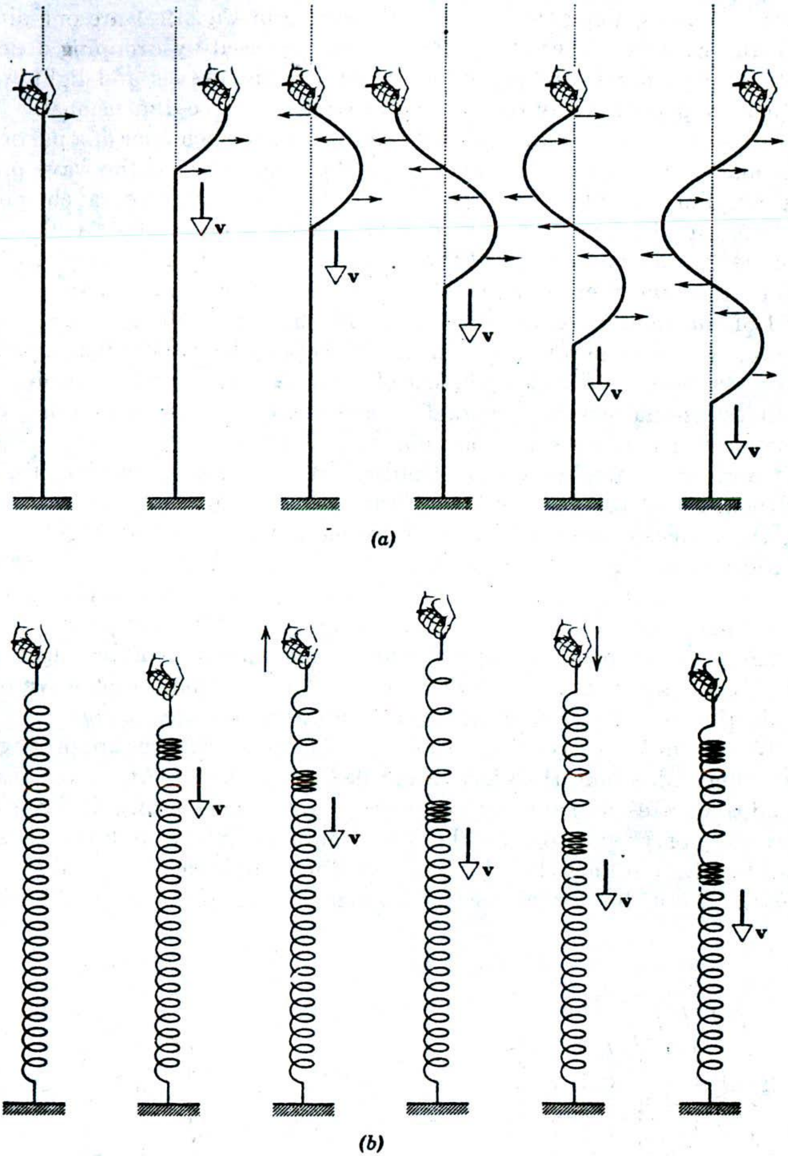


**Fig. 19-1** (*a*) In a transverse wave the particles of the medium (string) vibrate at right angles to the direction in which the wave itself is propagated. (*b*) In a longitudinal wave the particles of the medium (spring) vibrate in the same direction as that in which the wave itself is propagated.

according to the number of dimensions in which they propagate energy. Waves moving along the string or the spring of Fig. 19–1 are one-dimensional. Surface waves or ripples on water, caused by dropping a pebble into a quiet pond, are two-dimensional. Sound waves and light waves which emanate radially from a small source are three-dimensional.

Waves may be classified further according to the behavior of a particle of the matter conveying the wave during the course of time the wave propagates. For example, we can produce a *pulse* or a *single wave* traveling down a taut rope by applying a single sidewise movement at its end. Each particle remains at rest until the pulse reaches it, then it moves during a short time, and then it again remains at rest. If we continue to move the end of the rope back and forth (Fig. 19–1*a*), we produce a *train of waves* traveling along the rope. If our motion is periodic, we produce a *periodic train of waves* in which each particle of the rope has a periodic motion. The simplest special case of a periodic wave is a *simple harmonic wave* which gives each particle a simple harmonic motion.

Consider a three-dimensional pulse. We can draw a surface through all points undergoing a similar disturbance at a given instant. As time goes on, this surface moves along showing how the pulse propagates. We can draw similar surfaces for subsequent pulses. For a periodic wave we can generalize the idea by drawing in surfaces, all of whose points are in the same phase of motion. These surfaces are called *wavefronts.* If the medium is homogeneous and isotropic, the direction of propagation is always at right angles to the wavefront. A line normal to the wavefronts, indicating the direction of motion of the waves, is called a *ray.*

Wavefronts can have many shapes. If the disturbances are propagated in a single direction, the waves are called *plane waves.* At a given instant conditions are the same everywhere on any plane perpendicular to the direction of propagation. The wavefronts are plane and the rays are parallel straight lines (Fig. 19–2*a*). Another simple case is that of *spherical waves.* Here the disturbance is propagated out in all directions from a

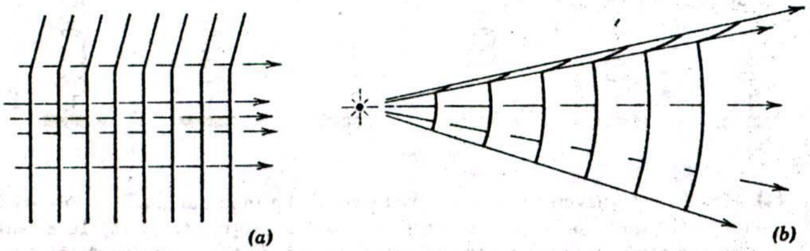


**Fig. 9–2** (*a*) A plane wave. The planes represent wavefronts spaced a wavelength apart, and the arrows represent rays. (*b*) A spherical wave. The rays are radial and the wavefronts, spaced a wavelength apart, form spherical shells. Far out from the source, however, small portions of the wavefronts become nearly plane.

point source of waves. The wavefronts are spheres and the rays are radial lines leaving the point source in all directions (Fig. 19–2*b*). Far from the source the spherical wavefronts have very small curvature, and over a limited region they can often be regarded as plane. Of course, there are many other possible shapes for wavefronts.

We shall refer to all these wave types as we progress through the wave phenomena of physics. In this chapter we often use the transverse wave in a string to illustrate the general properties of waves. In the next chapter we shall see the consequences of these properties for sound, a longitudinal mechanical wave. Later in the text we will discuss the properties of nonmechanical waves such as light and matter waves.

## 19–3 Traveling Waves

Let us consider a long string stretched in the $x$-direction along which a transverse wave is traveling. At some instant of time, say $t = 0$, the shape of the string can be represented by

$$y = f(x) \qquad t = 0, \tag{19–1}$$

where $y$ is the transverse displacement of the string at the position $x$. In Fig. 19–3*a* we show a possible waveform (a pulse) on the string at $t = 0$. Experiment shows that as time goes on such a wave travels along the string without changing its form, provided internal frictional losses are small enough. At some time $t$ later the wave has traveled a distance $vt$ to the right, where $v$ is magnitude of the wave velocity, assumed constant. The equation of the curve at the time $t$ is therefore

$$y = f(x - vt) \qquad t = t. \tag{19–2}$$

This gives us the same waveform about the point $x = vt$ at time $t$ as we had about $x = 0$ at the time $t = 0$ (Fig. 19–3*b*). Equation 19–2 is the

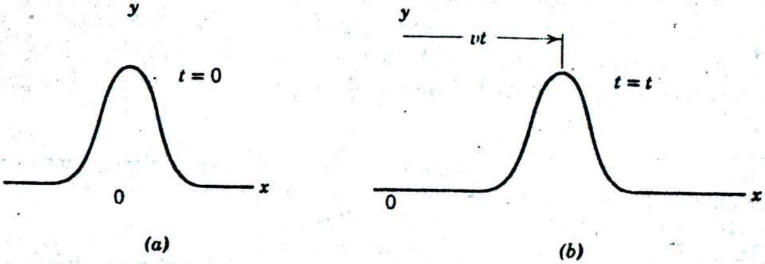


**Fig. 19–3** (*a*) The shape of a string (in this case a pulse) at $t = 0$. (*b*) At a later time $t$ the pulse has traveled to the right a distance $x = vt$.

general equation representing a wave of *any* shape *traveling to the right.* To describe a particular shape we must specify exactly what the function $f$ is.*

Let us look more carefully at this equation. If we wish to follow a particular part (or phase) of the wave as times goes on, then in the equation we look at a particular value of $y$ (say, the top of the pulse just described). Mathematically this means we look at how $x$ changes with $t$ when $(x - vt)$ has some particular fixed value. We see at once that as $t$ increases $x$ must increase in order to keep $(x - vt)$ fixed. Hence, Eq. 19–2 does in fact represent a wave traveling to the right (increasing $x$ as time goes on). If we wished to represent a wave *traveling to the left*, we would write

$$y = f(x + vt), \tag{19–3}$$

for here the position $x$ of some fixed phase $(x + vt)$ of the wave decreases as time goes on. The velocity of a particular phase of the wave is easily obtained. For a particular phase of a wave traveling to the right we require that

$$x - vt = \text{constant}.$$

Then differentiation with respect to time gives

$$\frac{dx}{dt} - v = 0 \qquad \text{or} \qquad \frac{dx}{dt} = v, \tag{19–4}$$

so that $v$ is really the *phase velocity* of the wave. For a wave traveling to the left we obtain $-v$, in the same way, as its phase velocity.†

The general equation of a wave can be interpreted further. Note that for any fixed value of the time $t$ the equation gives $y$ as a function of $x$. This defines a curve, and this curve represents the actual shape of the string at this chosen time. It gives us a snapshot of the wave at this time. Suppose, on the other hand, we wish to focus our attention on one point of the string, that is, a fixed value of $x$. Then the equation gives us $y$ as a function of the time $t$. This describes how the transverse position of this point on the string changes with time.

The argument just presented holds for longitudinal waves as well as for transverse waves. The analogous longitudinal example is that of a long straight tube of gas whose axis is taken as the $x$-axis, and the wave or pulse is a pressure change traveling along the tube. Then the same reasoning leads us to an equation, having the form of Eqs. 19–2 and 19–3, which gives the pressure variations with time at all points of the tube. (See Section 20–3.)

---

* When we say that "$y$ is a function of $(x - vt)$," we mean that the variables $x$ and $t$ occur only in the combination $x - vt$. For example, $\sin k(x - vt)$, $\log (x - vt)$, and $(x - vt)^3$ are functions of $x - vt$, but $x^2 - vt^2$ is not.

† In disturbances that can be represented as a group of waves, the energy may be transported with a velocity different from the phase velocity of any individual wave. This group velocity will be considered in Chapter 39 in connection with electromagnetic waves. Until then whenever we use the term wave velocity we mean the phase velocity of the wave.

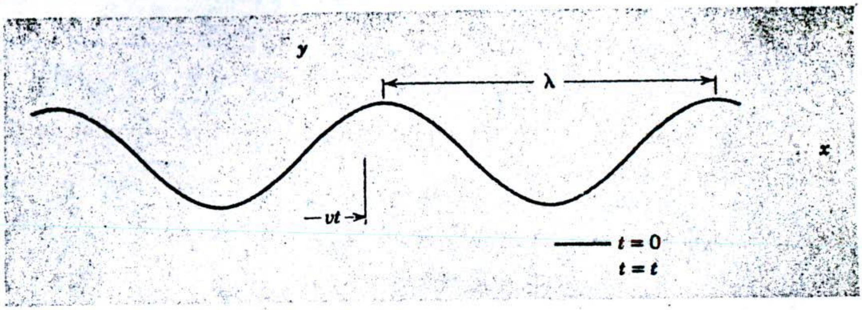


**Fig. 19-4** At $t = 0$, the string has a shape $y = y_m \sin 2\pi x/\lambda$ (solid line). At a later time $t$ the sine wave has moved to the right a distance $x = vt$, and the string has a shape given by $y = y_m \sin 2\pi(x - vt)/\lambda$.

Let us now consider a particular waveform, whose importance will soon become clear. Suppose that at the time $t = 0$ we have a wavetrain along the string given by

$$y = y_m \sin \frac{2\pi}{\lambda} x. \tag{19-5}$$

The wave shape is a sine curve (Fig. 19-4). The maximum displacement $y_m$ is the *amplitude* of the sine curve. The value of the transverse displacement $y$ is the same at $x$ as it is at $x + \lambda$, $x + 2\lambda$, etc. The symbol $\lambda$ is called the *wavelength* of the wavetrain and represents the distance between two adjacent points in the wave having the same phase. As times goes on let the wave travel to the right with a phase velocity $v$. Hence, the equation of the wave at the time $t$ is

$$y = y_m \sin \frac{2\pi}{\lambda} (x - vt). \tag{19-6}$$

Notice that this has the form required for a traveling wave (Eq. 19-2).

The *period* $T$ is the time required for the wave to travel a distance of one wavelength $\lambda$, so that

$$\lambda = vT. \tag{19-7}$$

Putting this relation into the equation of the wave, we obtain

$$y = y_m \sin 2\pi \left(\frac{x}{\lambda} - \frac{t}{T}\right). \tag{19-8}$$

From this form it is clear that $y$, at any given time, has the same value at $x + \lambda$, $x + 2\lambda$, etc., as it does at $x$, and that $y$, at any given position, has the same value at the time $t + T$, $t + 2T$, etc., as it does at the time $t$.

To reduce Eq. 19-8 to a more compact form, we define two quantities, the *wave number* $k$ and the *angular frequency* $\omega$ (see Eq. 15-12). They are given by

$$k = \frac{2\pi}{\lambda} \qquad \text{and} \qquad \omega = \frac{2\pi}{T}. \tag{19–9}$$

In terms of these quantities, the equation of a sine wave traveling to the right is

$$y = y_m \sin(kx - \omega t). \tag{19–10a}$$

For a sine wave traveling to the left, we have

$$y = y_m \sin(kx + \omega t). \tag{19–10b}$$

Comparing Eqs. 19–7 and 19–9, we see that the phase velocity $v$ of the wave is given by

$$v = \frac{\lambda}{T} = \frac{\omega}{k}. \tag{19–11}$$

In the traveling waves of Eqs. 19–10$a$ and 19–10$b$ we have assumed that the displacement $y$ is zero at the position $x = 0$ at the time $t = 0$. This, of course, need not be the case. The general expression for a sinusoidal wavetrain traveling to the right is

$$y = y_m \sin(kx - \omega t - \phi),$$

where $\phi$ is called the phase constant. For example, if $\phi = -90°$, the displacement $y$ at $x = 0$ and $t = 0$ is $y_m$. This particular example is

$$y = y_m \cos(kx - \omega t),$$

for the cosine function is displaced by 90° from the sine function.

If we fix our attention on a given point of the string, say $x = \pi/k$, the displacement $y$ at that point can be written* as

$$y = y_m \sin(\omega t + \phi).$$

This is similar to Eq. 15–29 for simple harmonic motion. Hence, any particular element of the string undergoes simple harmonic motion about its equilibrium position as this wavetrain travels along the string.

## 19–4 The Superposition Principle

It is an experimental fact that for many kinds of waves *two or more waves can traverse the same space independently of one another.* The fact that waves act independently of one another means that the displacement of any particle at a given time is simply the sum of the displacements that the individual waves alone would give it. This process of vector addition of the displacements of a particle is called *superposition.* For example, radio waves of many frequencies pass through a radio antenna; the electric currents set up in the antenna by the superposed action of all these waves are very complex. Nevertheless, we can still tune to a particular station,

---

* Using the fact that $\sin(\pi - \theta) = \sin\theta$.

the signal that we receive from it being in principle the same as that which we would receive if all other stations were to stop broadcasting. Likewise, in sound we can listen to notes played by individual instruments in an orchestra, even though the sound wave reaching our ears from the full orchestra is very complex.

For waves in deformable media the superposition principle holds whenever the mathematical relation between the deformation and the restoring force is one of simple proportionality. Such a relation is expressed mathematically by a linear equation. For electromagnetic waves the superposition principle holds because the mathematical relations between the electric and magnetic fields are linear.

The superposition principle seems so obvious that it is worthwhile to point out that it does not always hold. Superposition fails when the equations governing wave motion are not linear. Physically this happens when the wave disturbance is relatively large and the ordinary linear laws of mechanical action no longer hold. For example, beyond the elastic limit Hooke's law no longer holds and the linear relation $F = -kx$ can no longer be used.

As for sound, violent explosions create shock waves. Although shock waves are longitudinal elastic waves in air, they behave differently from ordinary sound waves. The equation governing their propagation is quadratic, and superposition does not hold. With two very loud notes the ear hears something more than just the two individual notes. Those familiar with high-fidelity apparatus will know that "intermodulation distortion" between two tones arises when the system fails to combine the tones linearly, and that this distortion is more apparent when the amplitude of the tones is high. A more obvious physical example is water waves. Ripples cannot travel independently across breakers as they can across gentle swells.

The importance of the superposition principle physically is that, where it holds, it makes it possible to analyze a complicated wave motion as a combination of simple waves. In fact, as was shown by the French mathematician J. Fourier (1768–1830), all that we need to build up the most general form of periodic wave are simple harmonic waves.* Fourier showed that any periodic motion of a particle can be represented as a combination of simple harmonic motions. For example, if $y(t)$ represents the motion of a source of waves having a period $T$, we can analyze $y(t)$ as follows:

$$\begin{aligned} y(t) = A_0 + A_1 \sin \omega t + A_2 \sin 2\omega t + A_3 \sin 3\omega t + \cdots \\ + B_1 \cos \omega t + B_2 \cos 2\omega t + B_3 \cos 3\omega t + \cdots \end{aligned}$$

where $\omega = 2\pi/T$. This expression is called a Fourier series. The $A$'s and $B$'s are constants which have definite values for any particular periodic motion $y(t)$. (See Fig. 19–5, for example.) If the motion is not periodic, as a pulse, the sum is replaced by an integral, the so-called Fourier integral. Hence, any motion of a source of waves can be represented in terms of simple harmonic motions. Since the motion of the source creates the

---

* See, for example, Thomas, *Calculus and Analytic Geometry*, Addison-Wesley, second edition, 1953, pp. 596-599.

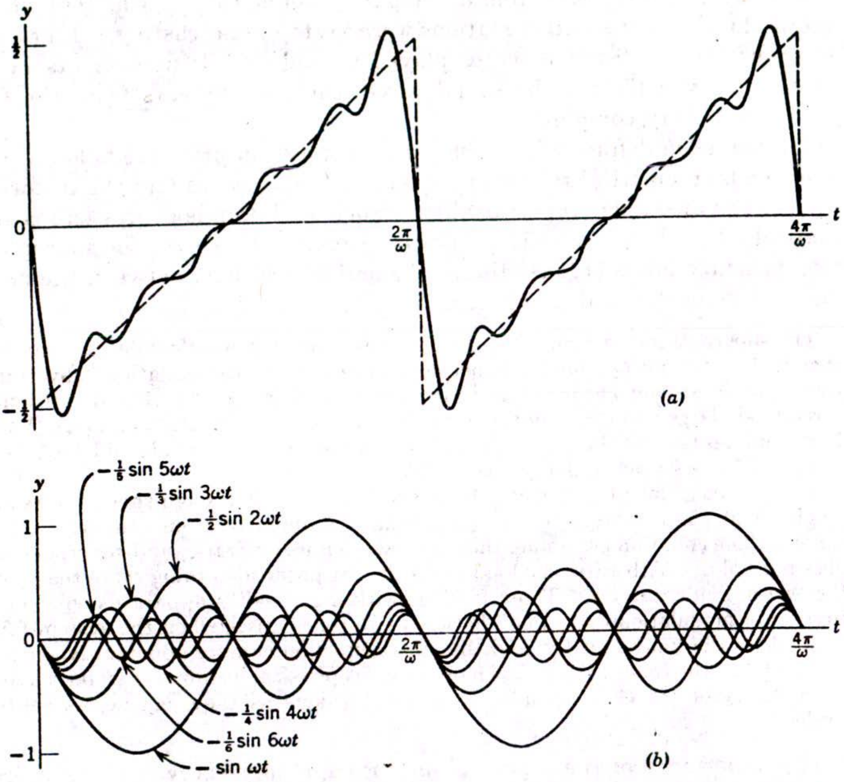


**Fig. 19–5** (*a*) The dashed line is a sawtooth "wave" commonly encountered in electronics. It can be written $y(t) = (\omega/2\pi)t - \frac{1}{2}$ for $0 < t < 2\pi/\omega$, as $y(t) = (\omega/2\pi)t - \frac{3}{2}$ for $2\pi/\omega < t < 4\pi/\omega$, etc. The Fourier series for this function is $y(t) = -\sin \omega t - \frac{1}{2} \sin 2\omega t - \frac{1}{3} \sin 3\omega t - \cdots$. The solid line is the sum of the first six terms of this series and can be seen to approximate the sawtooth quite closely, except for overshooting near the discontinuities. As more terms of the series are included, the approximation becomes better and better. (*b*) Here are shown the first six terms of the Fourier series which, when added together, yield the solid curve in (*a*).

waves, it should come as no surprise that the waves themselves can be analyzed as combinations of simple harmonic waves. Herein lies the importance of simple harmonic motion and simple harmonic waves.

## 19–5 Wave Speed

Given the characteristics of the medium it should be possible to calculate the wave speed from the basic principles of Newtonian mechanics. In this section we continue to focus our attention on transverse waves in a string and in Supplementary Topic III we show how to calculate the speed of such waves in the most general way. Here we consider two other approaches—a treatment based on dimensional analysis and a somewhat less general mechanical analysis in which we compute the speed of a transverse pulse along a stretched string.

We stated in Section 19–1 that the wave speed for a medium depends on the elasticity of the medium and on its inertia. For a stretched string the elasticity is measured by the tension $F$ in the string; the greater the tension the greater will be the elastic restoring force on an element of the string that is pulled sideways. The inertia characteristic is measured by $\mu$, the mass per unit length of the string. Assuming, then, that the wave speed $v$ depends only on $F$ and $\mu$, we can use dimensional analysis to find how $v$ depends on these quantities. In terms of mass $M$, length $L$, and time $T$, the dimensions of $F$ are $MLT^{-2}$ and the dimensions of $\mu$ are $ML^{-1}$. The only way these dimensions can be combined to get a velocity (which has the dimensions $LT^{-1}$) is to take the square root of $F/\mu$. That is, $F/\mu$ has the dimensions $L^2T^{-2}$ and $\sqrt{F/\mu}$ has the dimensions $LT^{-1}$ of a velocity. Dimensional analysis cannot account for any dimensionless quantities, so that the result

$$v = \sqrt{\frac{F}{\mu}} \tag{19–12}$$

may or may not be complete. The most we can say is that the wave speed is equal to a dimensionless constant times $\sqrt{F/\mu}$. The value of the constant can be obtained from a mechanical analysis of the problem or from experiment. These methods show that the constant is equal to unity and that Eq. 19–12 is correct as it stands.

Now let us *derive* the velocity of a pulse in a stretched string by a mechanical analysis. In Fig. 19–6 we show a wave pulse proceeding from right to left in the string with a speed $v$. We can imagine the entire string to be moved from left to right with this same speed so that the wave pulse remains fixed in space, whereas the particles composing the string successively pass through the pulse. This simply means that, instead of taking our reference frame to be the walls between which the string is stretched, we choose a reference frame which is in uniform motion with respect to that one. Because Newton's laws involve only accelerations, which are the same in both frames, we can use them in either frame. We just happen to choose a more convenient frame.

We consider a small section of the pulse of length $\Delta l$ to form an arc of a circle of radius $R$, as shown in the diagram. If $\mu$ is the mass per unit length of the string, the so-called linear density, then $\mu\,\Delta l$ is the mass of this element. The tension $F$ in the string is a tangential pull at each end of this small segment of the string. The horizontal components cancel and the vertical components are each equal to

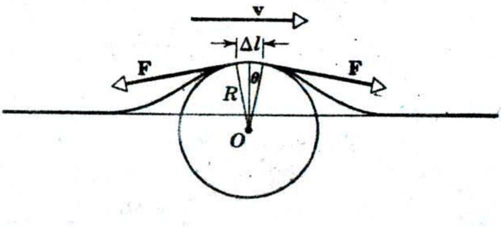


**Fig. 19–6** Derivation of wave speed by considering the forces on a section of string of length $\Delta l$.

F—32

$F \sin \theta$. Hence, the total vertical force is $2F \sin \theta$. Since $\theta$ is small, we can take $\sin \theta \cong \theta$ and

$$2F \sin \theta = 2F\theta = 2F \frac{(\Delta l/2)}{R} = F \frac{\Delta l}{R}.$$

This gives the force supplying the centripetal acceleration of the string particles directed toward $O$. Now the centripetal force acting on a mass $\mu \Delta l$ moving in a circle of radius $R$ with speed $v$ is $\mu \Delta l\, v^2/R$; see Section 6–3. Notice that the tangential velocity $v$ of this mass element along the top of the arc is horizontal and is the same as the pulse phase velocity. Combining the equivalent expressions just given we obtain

$$F \frac{\Delta l}{R} = \frac{\mu \Delta l\, v^2}{R}$$

or

$$v = \sqrt{\frac{F}{\mu}}.$$

If the amplitude of the pulse were very large compared to the length of the string, we would not have been able to use the approximation $\sin \theta \cong \theta$. Furthermore, the tension $F$ in the string would be changed by the presence of the pulse, whereas we assumed $F$ to be unchanged from the original tension in the stretched string. Therefore, our result, like superposition, holds only for relatively small transverse displacements of the string—which case, however, is widely applicable in practice. Notice also that the wave speed is independent of the shape of the wave, for no particular assumption about the actual shape of the pulse was used in the proof.

The frequency of a wave is naturally determined by the frequency of the source. The speed with which the wave travels through a medium is determined by the properties of the medium, as illustrated before. Once the frequency $\nu$ and speed $v$ of the wave are determined, the wavelength $\lambda$ is fixed. In fact, from Eq. 19–7 and the relation, $\nu = 1/T$, we have

$$\lambda = \frac{v}{\nu}. \tag{19–13}$$

▶ **Example 1.** A transverse sinusoidal wave is generated at one end of a long horizontal string by a bar which moves the end up and down through a distance of $\frac{1}{2}$ ft. The motion is continuous and is repeated regularly twice each second.

(*a*) If the string has a linear density of 0.0050 slug/ft and is kept under a tension of 2.0 lb, find the speed, amplitude, frequency, and wavelength of the wave motion.

The end moves $\frac{1}{4}$ ft away from the equilibrium position, first above it, then below it, for a total displacement of $\frac{1}{2}$ ft. Therefore, the amplitude $y_m$ is $\frac{1}{4}$ ft.

The entire motion is repeated twice each second so that the frequency is 2.0 vibrations per second.

The wave speed is given by $v = \sqrt{F/\mu}$. But $F = 2.0$ lb and $\mu = 0.0050$ slug/ft, so that

$$v = \sqrt{\frac{2.0 \text{ lb}}{0.0050 \text{ slug/ft}}} = 20 \text{ ft/sec}.$$

The wavelength is given by $\lambda = v/\nu$, so that

$$\lambda = \frac{20 \text{ ft/sec}}{2.0 \text{ vib/sec}} = 10 \text{ ft}.$$

(*b*) Assuming the wave moves from left to right and that, at $t = 0$, the end of the string described by $x = 0$ is in its equilibrium position $y = 0$, write the equation of the wave.

The general expression for a transverse sinusoidal wave moving from left to right is

$$y = y_m \sin (kx - \omega t - \phi).$$

Requiring that $y = 0$ for the conditions $x = 0$ and $t = 0$ yields

$$0 = y_m \sin (-\phi),$$

which means that the phase constant $\phi$ must be zero.* Hence for this wave

$$y = y_m \sin (kx - \omega t),$$

and with the values just found,

$$y_m = \tfrac{1}{4} \text{ ft} = 0.25 \text{ ft},$$

$$\lambda = 10 \text{ ft} \quad \text{or} \quad k = \frac{2\pi}{\lambda} = \frac{\pi}{5} \text{ ft}^{-1},$$

$$v = 20 \text{ ft/sec} \quad \text{or} \quad \omega = vk = 4\pi \text{ sec}^{-1},$$

we obtain as the equation for the wave

$$y = 0.25 \sin (0.2\pi x - 4\pi t),$$

where $x$ and $y$ are in feet and $t$ is in seconds.

**Example 2.** As this wave passes along the string, each particle of the string moves up and down at right angles to the direction of the wave motion. Find the velocity and acceleration of a particle 10 ft from the end.

The general form of this wave is

$$y = y_m \sin (kx - \omega t) = y_m \sin k(x - vt).$$

The $v$ in this equation is the constant horizontal velocity of the wavetrain. What we are after now is the velocity of a particle in the string through which this wave moves; this particle velocity is neither horizontal nor constant. In fact, each particle moves vertically, that is, in the $y$-direction. In order to determine the particle velocity, which we shall designate by the symbol $u$, let us fix our attention on a particle at a particular position $x$—that is, $x$ is now a constant in this equation—and ask how the particle displacement $y$ changes with time. With $x$ constant we obtain

$$u = \frac{\partial y}{\partial t} = -y_m \omega \cos (kx - \omega t),$$

in which the *partial derivative* notation $\partial y/\partial t$ reminds us that although in general $y$ is a function of both $x$ and $t$, we here assume that $x$ remains constant so that $t$ becomes the only variable. The acceleration $a$ of the particle at this (constant)

---

* It could also be $\pi$, $2\pi$, $3\pi$, etc., but these values for $\phi$ will not change our final results, as the student should show.

value of $x$ is

$$a = \frac{\partial^2 y}{\partial t^2} = \frac{\partial u}{\partial t} = -y_m\omega^2 \sin(kx - \omega t) = -\omega^2 y.$$

This shows that for each particle through which this transverse sinusoidal wave passes we have precisely SHM (simple harmonic motion), for the acceleration $a$ is proportional to the displacement $y$, but oppositely directed.

For a particle at $x = 10$ ft with the wave of Example 1, in which

$$y_m = 0.25 \text{ ft}, \qquad k = \frac{\pi}{5} \text{ ft}^{-1}, \qquad \omega = 4\pi \text{ sec}^{-1},$$

we obtain

$$u = -y_m\omega \cos(kx - \omega t)$$

or

$$u = -(0.25)\left[4\pi \cos\left(\frac{10\pi}{5} - 4\pi t\right)\right] = -\pi \cos(2\pi - 4\pi t),$$

and

$$a = -\omega^2 y$$

or

$$a = -(4\pi)^2 0.25 \sin(0.2\pi x - 4\pi t) = -4\pi^2 \sin(2\pi - 4\pi t),$$

where $t$ is expressed in seconds $u$ in ft/sec and $a$ in ft/sec$^2$.

Can you describe the motion of this particle at the time $t = 4$ sec? ◄

## 19–6 Power and Intensity in Wave Motion

In Fig. 19-7 we draw an element of the stretched string at some position $x$ and at a particular time $t$. The *transverse* component of the tension in the string exerted *by* the element to the left of $x$ *on* the element of the right of $x$ is

$$F_{\text{trans}} = -F\frac{\partial y}{\partial x}.$$

**F** is the tension in the string; $\partial y/\partial x$ gives the tangent of the angle made by the direction of **F** with the horizontal at the time $t$ in question and, because we assume small displacements, this can be taken equal also to the sine of the angle. The transverse force is in the direction of increasing $y$; in the figure the slope is negative, so the transverse force is positive. The transverse velocity of the particle at $x$ is $\partial y/\partial t$, which may be positive or negative. The power being expended by

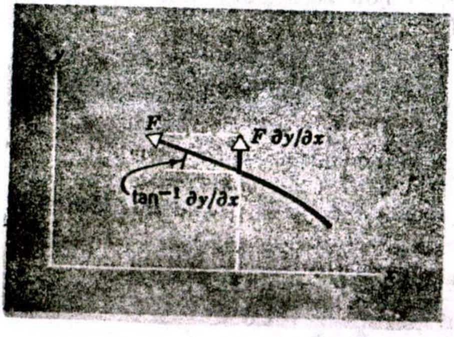


**Fig. 19–7** The transverse component of the tension in the string at each point $x$ is $F\ (\partial y/\partial x)$.

the force at $x$, or the energy passing through the position $x$ per unit time in the positive $x$-direction (see Section 7–6), is

$$P = \left(-F\frac{\partial y}{\partial x}\right)\frac{\partial y}{\partial t}.$$

Suppose that the wave on the string is the simple sine wave

$$y = y_m \sin(kx - \omega t).$$

Then the magnitude of the slope at $x$ is

$$\frac{\partial y}{\partial x} = ky_m \cos(kx - \omega t), \qquad [t = \text{constant}]$$

and the transverse force is

$$-F\frac{\partial y}{\partial x} = -Fky_m \cos(kx - \omega t).$$

The transverse velocity of a particle of the string at $x$ is

$$u = \frac{\partial y}{\partial t} = -\omega y_m \cos(kx - \omega t), \qquad [x = \text{constant}].$$

Hence, the power transmitted through $x$ is

$$\begin{aligned} P &= (-Fky_m)(-\omega y_m)\cos^2(kx - \omega t), \\ &= y_m{}^2 k\omega F \cos^2(kx - \omega t). \end{aligned}$$

Notice that the power or rate of flow of energy is not constant. The power is not constant because the power input oscillates. As the energy is passed along the string, it is stored in each element of string as a combination of kinetic energy of motion and potential energy of deformation. The situation is much like that in an alternating current circuit; there energy is stored both in the inductor and in the capacitor and the power input oscillates. For a string the power is absorbed by internal friction and viscous effects and appears as heat; in the circuit the power is expended in the resistor and appears as heat. The power input to the string or the circuit is often taken to be the *average* over one period of motion. The average power delivered is

$$\bar{P} = \frac{1}{T}\int_t^{t+T} P\,dt,$$

where $T$ is the period. Using the fact that the average value of $\sin^2 x$ or $\cos^2 x$ over one cycle is $\frac{1}{2}$, we obtain for the string

$$\bar{P} = \tfrac{1}{2}y_m{}^2 k\omega F = 2\pi^2 y_m{}^2 \nu^2 \frac{F}{v},$$

a result which does not depend on $x$ or $t$. For the string, however, $v = \sqrt{F/\mu}$, so that

$$\bar{P} = 2\pi^2 y_m{}^2 \nu^2 \mu v.$$

The fact that the rate of transfer of energy depends on the square of the wave amplitude and square of the wave frequency is true in general, holding for all types of waves.

The student should confirm the fact that, if we had picked a wave traveling in the negative $x$-direction, we would have obtained the negative of this result. That is, the wave delivers power in the direction of wave propagation.

In a three-dimensional wave, such as a light wave or a sound wave from a point source, it is more significant to speak of the *intensity* of the wave. Intensity is defined as the power transmitted across a unit area normal to the direction in which the wave is traveling. Just as with power in the wave in a string, the intensity of a space wave is always proportional to the square of the amplitude.

As a wave progresses through space, its energy may be absorbed. For example, in a viscous medium, such as syrup or lead, mechanical waves would rapidly decay in amplitude and disappear, owing to absorption of energy by internal friction. In most cases of interest to us, however, absorption will be negligible. Throughout this chapter we have assumed that there is no loss of energy in a given wave, no matter how far it travels.

▶ **Example 3.** Spherical waves travel from a source of waves whose power output is $P$; see Fig. 19–8. Find how the wave intensity depends on the distance from the source. We assume that the medium is isotropic and that the source radiates uniformly in all directions, that is, that its emission is spherically symmetrical.

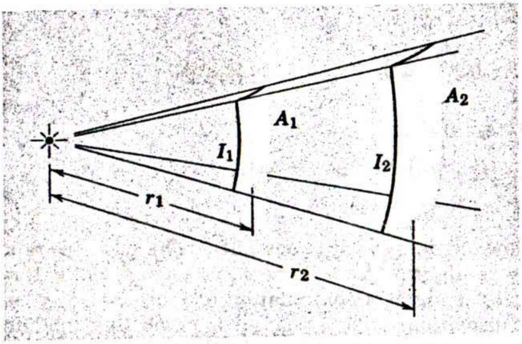


**Fig. 19–8** Example 3.

The intensity of a 3-dimensional wave is the power transmitted across a unit area normal to the direction of propagation. As the wavefront expands from a distance $r_1$ from the source at the center to a distance $r_2$, its surface area increases from $4\pi r_1^2$ to $4\pi r_2^2$. If there is no absorption of energy, the total energy transported per second by the wave remains constant at the value $P$, so that

$$P = 4\pi r_1^2 I_1 = 4\pi r_2^2 I_2,$$

where $I_1$ and $I_2$ are the wave intensities at $r_1$ and $r_2$ respectively. Hence,

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

and the wave intensity varies inversely as the square of its distance from the source. Since the intensity is proportional to the square of the amplitude, the amplitude of the wave must vary inversely as the distance from the source. ◀

## 19–7 Interference of Waves

*Interference* refers to the physical effects of superimposing two or more wave trains. Let us consider two waves of equal frequency and amplitude traveling with the same speed in the same direction ($+x$) but with a phase difference $\phi$ between them. The equations of the two waves will be

$$y_1 = y_m \sin(kx - \omega t - \phi) \tag{19-14}$$

and

$$y_2 = y_m \sin(kx - \omega t). \tag{19-15}$$

We can rewrite the first equation in two equivalent forms

$$y_1 = y_m \sin\left[k\left(x - \frac{\phi}{k}\right) - \omega t\right]. \tag{19-14a}$$

or

$$y_1 = y_m \sin\left[kx - \omega\left(t + \frac{\phi}{\omega}\right)\right]. \tag{19-14b}$$

Equations 19-14*a* and 19-15 suggest that if we take a "snapshot" of the two waves at any time $t$, we will find them displaced from one another along the $x$-axis by the constant distance $\phi/k$. Equations 19-14*b* and 19-15 suggest that if we station ourselves at any position $x$ the two waves will give rise to two simple harmonic motions having a constant time difference $\phi/\omega$. This gives some insight into the meaning of the phase difference $\phi$.

Now let us find the resultant wave, which, on the assumption that superposition occurs, is the sum of Eqs. 19-14 and 19-15 or

$$y = y_1 + y_2 = y_m[\sin(kx - \omega t - \phi) + \sin(kx - \omega t)].$$

From the trigonometric equation for the sum of the sines of two angles

$$\sin B + \sin C = 2 \sin \tfrac{1}{2}(B + C) \cos \tfrac{1}{2}(C - B), \tag{19-16}$$

we obtain

$$y = y_m\left[2 \sin\left(kx - \omega t - \frac{\phi}{2}\right) \cos\frac{\phi}{2}\right],$$

$$= \left(2y_m \cos\frac{\phi}{2}\right) \sin\left(kx - \omega t - \frac{\phi}{2}\right). \tag{19-17}$$

This resultant wave corresponds to a new wave having the same frequency but with an amplitude $2y_m \cos(\phi/2)$. If $\phi$ is *very small* (compared to 180°), the resultant amplitude will be nearly $2y_m$. That is, when $\phi$ is very small, $\cos(\phi/2) \cong \cos 0° = 1$. When $\phi$ is *zero*, the two waves have the same phase everywhere. The crest of one corresponds to the crest of the other and likewise for the troughs. The waves are then said to interfere constructively. The resultant amplitude is just twice that of either wave alone. If $\phi$ is *near* 180°, on the other hand, the resultant amplitude will be *nearly* zero. That is, when $\phi \cong 180°$, $\cos(\phi/2) \cong \cos 90° = 0$. When $\phi$ is *exactly* 180°, the crest of one wave corresponds exactly to the trough of the other. The waves are then said to interfere destructively. The resultant amplitude is zero.

In Fig. 19-9*a* we show the superposition of two wavetrains almost in phase ($\phi$ small) and in Fig. 19-9*b* the superposition of two wavetrains almost 180° out of phase ($\phi \cong 180°$). Notice that in these figures the

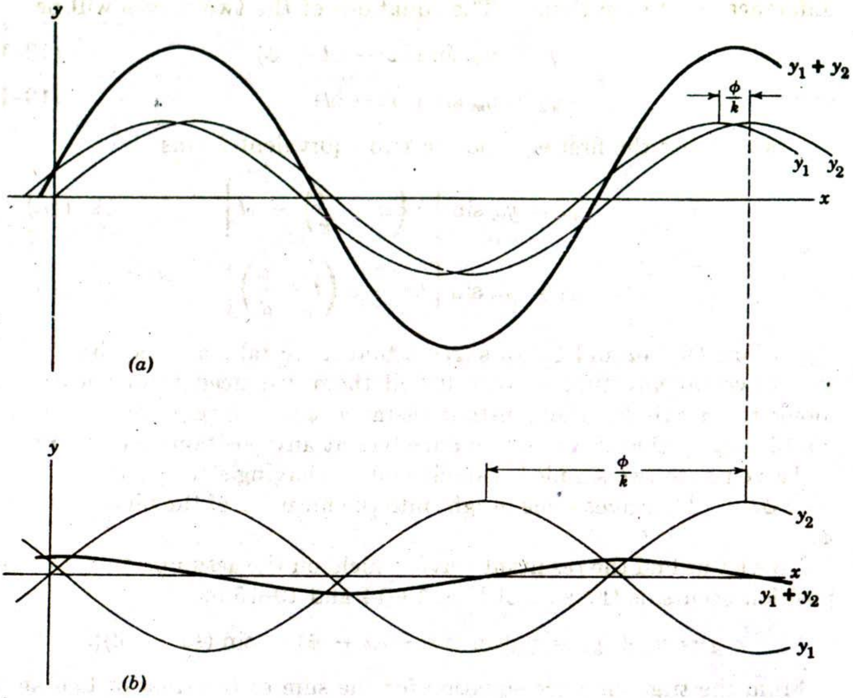


**Fig. 19–9** (*a*) The superposition of two waves of equal frequency and amplitude that are almost in phase results in a wave of almost twice the amplitude of either component. (*b*) The superposition of two waves of equal frequency and amplitude and almost 180° out of phase results in a wave whose amplitude is nearly zero. Note that in both the resultant frequency is unchanged. (The drawings correspond to the instant $t = 0$.)

algebraic sum of the ordinates of the thin (component) curves at any value of $x$ equals the ordinate of the thick (resultant) curve. The sum of two waves can, therefore, have different values, depending on their phase relations.

The resultant wave will be a sine wave, even when the amplitudes of the component sine waves are unequal. Figure 19–10, for example, illustrates the addition of two sine waves of the same frequency and velocity but different amplitudes. The resultant amplitude depends on the phase difference, which is taken as zero in this figure. The result for other phase differences could be obtained by shifting one of the component waves sideways with respect to the other and would give a smaller resultant amplitude. The smallest resultant amplitude would be the difference in the amplitudes of the components, obtained when the phases differ by 180°. However, the resultant is always a sine wave. The addition of any number of sine waves having the same frequency and velocity gives a

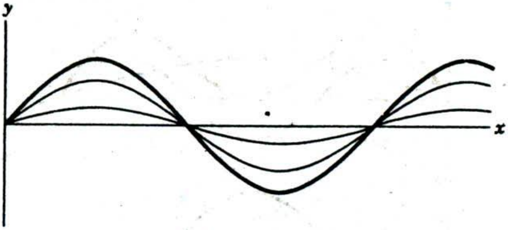


**Fig. 19-10** The addition of two waves of same frequency and phase but differing amplitudes (light lines) yields a third wave of the same frequency and phase (heavy line).

similar result. The resultant waveform will always have a constant amplitude because the component waves (and their resultant) all move with the same velocity and maintain the same relative position. The actual state of affairs can be pictured by having all the waves in Figs. 19-9 and 19-10 move toward the right with the same speed.

In practice, interference effects are obtained from wavetrains which originate in the same source (or in sources having a fixed phase relationship to one another) but which follow different paths to the point of interference. The phase difference $\phi$ between the waves arriving at a point can be calculated by finding the difference between the paths traversed by them from the source to the point of interference. The path difference is $\phi/k$ or $(\phi/2\pi)\lambda$. When the path difference is 0, $\lambda$, $2\lambda$, $3\lambda$, etc., so that $\phi = 0$, $2\pi$, $4\pi$, etc., the two waves interfere constructively. For path differences of $\frac{1}{2}\lambda$, $\frac{3}{2}\lambda$, $\frac{5}{2}\lambda$, etc., $\phi$ is $\pi$, $3\pi$, $5\pi$, etc., and the waves interfere destructively. We shall return to these matters later in more detail.

## 19-8 Complex Waves

The waves we have considered so far have been of the simple harmonic type, in which the displacements at any time are represented by a sine curve. We have seen that superposition of any number of such waves having the same frequency and velocity, but arbitrary amplitudes and phases, still gives rise to a resultant wave of this simple type. If, however, we superimpose waves that have *different frequencies*, the resulting wave is *complex*. In a complex wave the motion of a particle is no longer simple harmonic motion, and the wave shape is no longer a sine curve. In this section we consider only the qualitative aspects of complex waves. The analytical treatment of such waves will be given when we encounter physical situations described by them. We will look at the results of adding graphically two or more waves traveling with the same speed in the same direction but having various relative frequencies, amplitudes, and phases.

In Figs. 19-11*a* and 19-11*b* we add two waves having the same amplitude but having frequencies in the ratio 3 to 1; the phase relation is changed from *a* to *b* and we see how changing the phase relation may produce a resultant of very different form. If these represent sound waves, our eardrums will vibrate in a way represented by the resultant in each case, but we will hear and interpret these as the two original frequencies, regardless of their phase relation. If the resultant waves represent visible light, our eyes will receive the same sensation of a mixture of two colors, regardless of the phase relation of the components.

In Fig. 19-12 three waves of different frequencies, and amplitudes are added. The resultant complex wave is quite different from a simple periodic wave and, in this respect resembles waveforms normally generated by musical instruments. In Fig. 19-13 a wave of very high frequency is added to one of very low frequency.

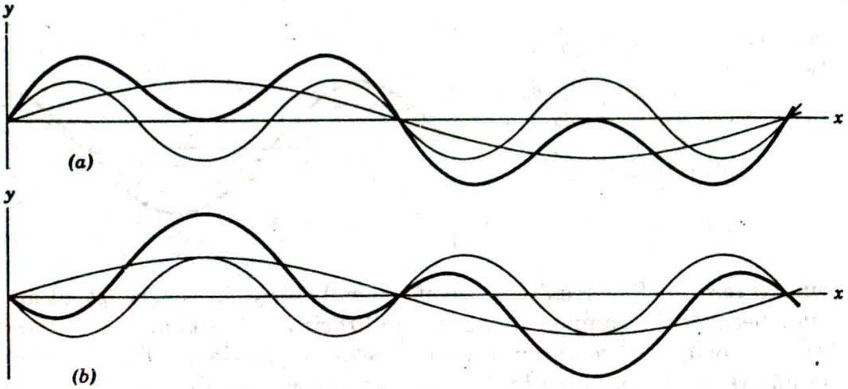


**Fig. 19–11** The addition of two waves with frequency ratio 3:1 (light lines) yields a wave whose shape (heavy line) depends on the phase relationship of the components Compare (*a*) and (*b*).

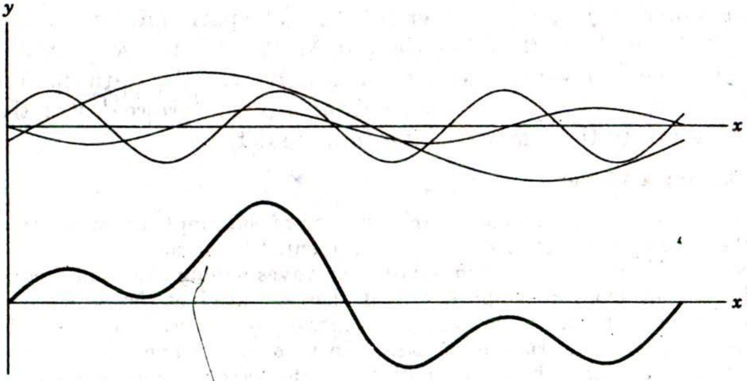


**Fig. 19–12** The addition of three waves (top) of differing frequencies yields a complex waveform (bottom).

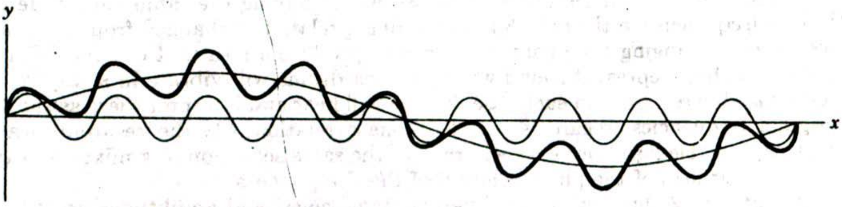


**Fig. 19–13** The addition (heavy line) of two waves of widely differing frequency (light lines).

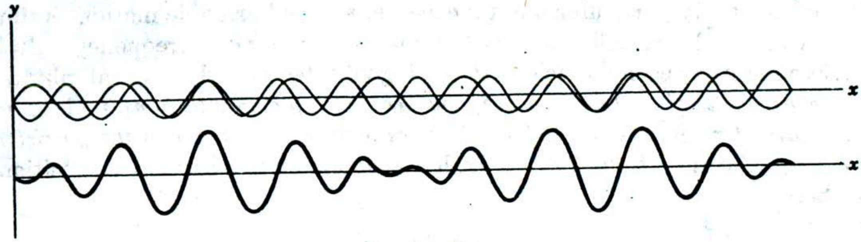


**Fig. 19–14** The addition (bottom) of two waves with nearly the same frequency (top), illustrating the phenomenon of beats. (see Chapter 20.)

Each component frequency is clearly discernible in the resultant. In Fig. 19–14 two waves of nearly the same frequency are added. The resultant wave consists of groups which, in the case of sound, produce the familiar phenomenon of beats (Section 20–6).

In all of these figures the resultant wave is obtained under the assumption that the principle of superposition holds, by simply adding the displacements caused by the individual waves at every point. Because all the component waves travel with the same velocity, the resultant waveform moves with this same velocity and the wave shape is unchanged.

The cathode-ray oscilloscope (Chapter 27) gives the simplest way of observing how complex waves can be synthesized and analyzed in terms of simple harmonic waves.

## 19–9 Standing Waves

In a one-dimensional body of finite size, such as a taut string held by two clamps a distance $l$ apart, traveling waves in the string are reflected from the boundaries of the body, that is, from the clamps. Each such reflection gives rise to a wave traveling in the string in the opposite direction. The reflected waves add to the incident waves according to the principle of superposition.

Consider two wavetrains of the same frequency, speed, and amplitude which are traveling in *opposite directions* along a string. Two such waves may be represented by the equations

$$y_1 = y_m \sin (kx - \omega t),$$

$$y_2 = y_m \sin (kx + \omega t).$$

Hence, the resultant may be written as

$$y = y_1 + y_2 = y_m \sin (kx - \omega t) + y_m \sin (kx + \omega t) \qquad (19\text{–}18a)$$

or, making use of the trigonometric relation of Eq. 19–16, as

$$y = 2y_m \sin kx \cos \omega t. \qquad (19\text{–}18b)$$

Equation 19–18*b* is the equation of a *standing* wave.* Notice that a

---

* Standing waves may also be produced in finite bodies of two or three dimensions; see Sections 20–5 and 38–5 respectively for examples.

particle at any particular point $x$ executes simple harmonic motion as time goes on, and that all particles vibrate with the same frequency. In a traveling wave each particle of the string vibrates with the same amplitude. *Characteristic of a standing wave, however, is the fact that the amplitude is not the same for different particles but varies with the location $x$ of the particle.** In fact, the amplitude, $2y_m \sin kx$, has a *maximum* value of $2y_m$ at positions where

$$kx = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \text{ etc.}$$

or

$$x = \frac{\lambda}{4}, \frac{3\lambda}{4}, \frac{5\lambda}{4}, \text{ etc.}$$

These points are called *antinodes* and are spaced one-half wavelength apart. The amplitude has a *minimum* value of zero at positions where

$$kx = \pi, 2\pi, 3\pi, \text{ etc.}$$

or

$$x = \frac{\lambda}{2}, \lambda, \frac{3\lambda}{2}, 2\lambda, \text{ etc.}$$

These points are called *nodes* and are spaced one-half wavelength apart. The separation between a node and an adjacent antinode is one-quarter wavelength.

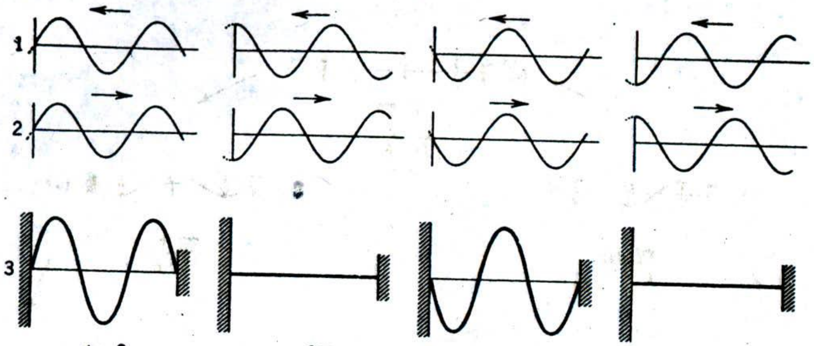
It is clear that energy is not transported along the string to the right or to the left, for energy cannot flow past the nodal points in the string which are permanently at rest. Hence, the energy remains "standing" in the string, although it alternates between vibrational kinetic energy and elastic potential energy. We call the motion a wave motion because we can think of it as a superposition of waves traveling in opposite directions (Eq. 19–18*a*). We can equally well regard the motion as an oscillation of the string as a whole (Eq. 19–18*b*), each particle oscillating with SHM of angular frequency $\omega$ and with an amplitude that depends on its location. Each small part of the string has inertia and elasticity, and the string as a whole can be thought of as a collection of coupled oscillators. Hence, the vibrating string is the same in principle† as a spring-mass system, except that a spring-mass system has only one natural frequency, and a vibrating string has a large number of natural frequencies (Section 19–10).

In Fig. 19–15, in (*a*), (*b*), (*c*), and (*d*), we show a standing wave pattern separately at intervals of one-quarter of a period in the lower figures, 3. The traveling waves, one moving in the positive $x$-direction and the other moving in the negative $x$-direction, whose superposition can be considered

---

* The combining waves moving in opposite directions along the string will still produce standing waves even if their amplitudes are unequal. We consider only the equal-amplitude case here; see Problem 25, however.

† For a general discussion see "On the Teaching of 'Standing Waves,'" J. Rekveld, *American Journal of Physics*, March 1958, p. 159.



$t = 0$ (a) $t = \frac{1}{4}T$ (b) $t = \frac{1}{2}T$ (c) $t = \frac{3}{4}T$ (d)

**Fig. 19-15** Standing waves as the superposition of left- and right-going waves; 1 and 2 are the components, 3 the resultant.

to give rise to the standing wave, are shown for the same quarter-period intervals in the upper figures 2 and 1. Standing waves can also be produced with electromagnetic waves and with sound waves.

In Fig. 19-16 we show how the energy associated with the oscillating string shifts back and forth between kinetic energy of motion $K$ and potential energy of deformation $U$ during one cycle. The student should compare this with Fig. 8-4, which shows the same thing for a mass-spring oscillator. Oscillating strings often vibrate so rapidly that the eye perceives only a blur whose shape is that of the envelope of the motion; see Fig. 19-17.

The superposition of an incident wave and a reflected wave, being the sum of two waves traveling in opposite directions, will give rise to a standing wave. We shall now consider the process of reflection of a wave more closely. Suppose a pulse travels down a stretched string which is fixed at one end, as shown in Fig. 19-18$a$. When the pulse arrives at that end, it exerts an upward force on the support. The support is rigid, however, and does not move. By Newton's third law the support exerts an equal but oppositely directed force on the string. This reaction force generates a pulse at the support, which travels back along the string in a direction opposite to that of the incident pulse. We say that the incident pulse has been *reflected* at the fixed end point of the string. Notice that the reflected pulse returns with its transverse displacement reversed. If a wavetrain is incident on the fixed end point, a reflected wavetrain is generated at that point in the same way. The displacement of any point along the string is the sum of the displacements caused by the incident and reflected wave. Since the end point is fixed, these two waves must always interfere destructively at that point so as to give zero displacement there. Hence, the reflected wave is always 180° out of phase with the incident wave at a fixed

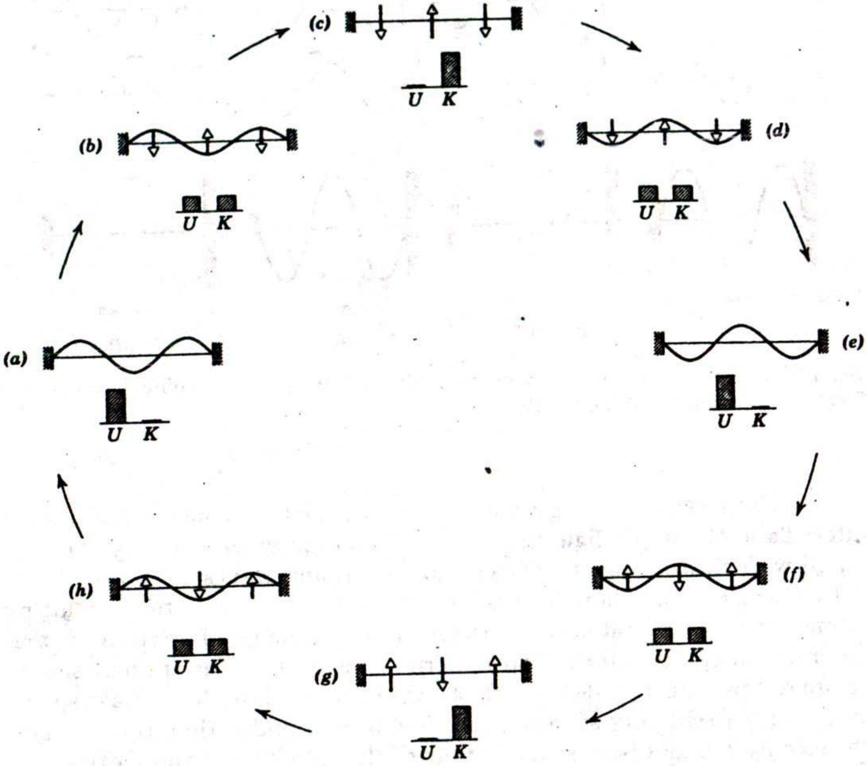


**Fig. 19–16** A standing wave in a stretched string, showing one cycle of oscillation. At (*a*) the string is momentarily at rest and the energy of the system is all potential energy of elastic deformation associated with the transverse displacement of the string. (*b*) An eighth-cycle later the displacement is reduced and the string is in motion. The two arrows show the velocities of the string particles at the positions shown. *K* and *U* have the same value. (*c*) The string is not displaced, but its particles have their maximum speeds; the energy is all kinetic. The motion continues until the initial condition (*a*) is reached after which the cycle continues to repeat itself.

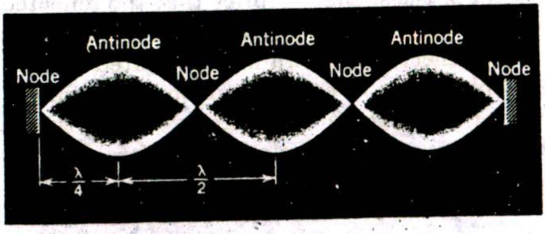


**Fig. 19–17** The envelope of a standing wave, corresponding to a time exposure of the motion, and showing the patterns of nodes and antinodes.

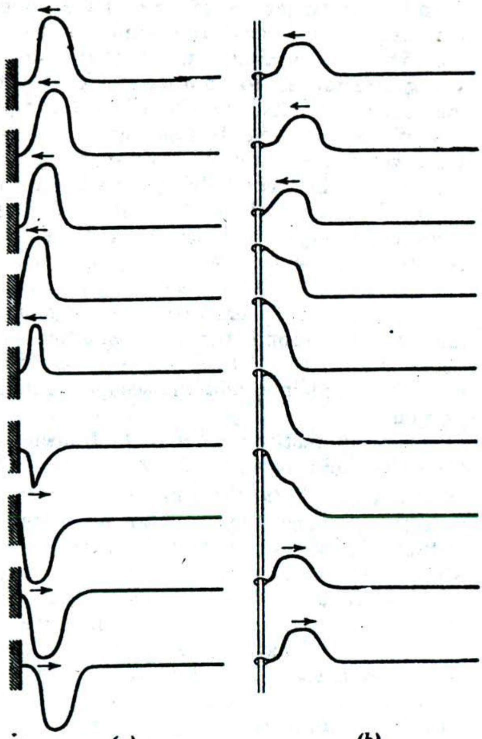


**Fig. 19–18** (*a*) Reflection of a pulse at the fixed end of a string. The drawings are spaced uniformly in time. The phase is changed by 180° on reflection. (*b*) Reflection of a pulse at an end free to move in a transverse direction. (The string is attached to a ring which slides vertically without friction.) The phase is unchanged on reflection.

boundary. We say that *on reflection from a fixed end a wave undergoes a phase change of 180°.*

Let us now consider the reflection of a pulse at a free end of a stretched string, that is, at an end that is free to move transversely. This can be achieved by attaching the end to a very light ring free to slide without friction along a transverse rod, or (see later) to a long and very much lighter string. When the pulse arrives at the free end, it exerts a force on the element of string there. This element is accelerated and its inertia carries it past the equilibrium point; it "overshoots" and exerts a reaction force on the string. This generates a pulse which travels back along the string in a direction opposite to that of the incident pulse. Once again we get reflection, but now at a free end. The free end will obviously suffer the maximum displacement of the particles on the string; an incident and a reflected wavetrain must interfere constructively at that point if we are to have a maximum there. Hence, the reflected wave is always in phase with the incident wave at that point (see Fig. 19–18*b*). We say that *at a free end a wave is reflected without change of phase.*

Hence, when we have a standing wave in a string, there will be a node at a fixed end and an antinode at a free end. These ideas will be applied to sound waves and electromagnetic waves in subsequent chapters.

In the treatment just given we have assumed that there is total reflection at the boundary. In general, at a boundary there is partial reflection and partial transmission. For example, suppose that instead of being attached to a rigid wall the string is attached to another string. At the boundary joining the strings the incident wave will be partly reflected and partly transmitted. The amplitude of the reflected wave will be less than that of the incident wave because a transmitted wave continues along the second string and carries away some of the incident energy. If the second string has a greater linear density than the first, the wave reflected back into the first will still suffer a phase shift of 180° on reflection. But because its amplitude is less than the incident wave, the boundary point will not be a node and will move. Thus a net energy transfer occurs along the first string into the second. If the second string has a smaller linear density than the first, partial reflection occurs without change of phase, but once again energy is transmitted to the second string. In practice the best way to realize a "free end" for a string is to attach it to a long and very much lighter string. The energy transmitted is negligible, and the second string serves to maintain the tension in the first one.

It is of interest to note that the transmitted wave travels with a different speed than the incident and reflected waves. The wave speed is determined by the relation $v = \sqrt{F/\mu}$; the tension is the same in both strings, but their densities are different. Hence, the wave travels more slowly in the denser string. The frequency of the transmitted wave is the same as that of the incident and reflected waves. Waves having the same frequency but traveling with different speeds have different wavelengths. Hence, from the relation $\lambda = v/\nu$ we conclude that in the denser string, where $v$ is less, the wavelength is shorter. This phenomenon of change of wavelength as a wave passes from one medium to another will be encountered frequently in our study of light waves.

## 19–10 Resonance

In general, whenever a system capable of oscillating is acted on by a periodic series of impulses having a frequency equal or nearly equal to one of the natural frequencies of oscillation of the system, the system is set into oscillation with a relatively large amplitude. This phenomenon is called *resonance* (see Section 15–10) and the system is said to resonate with the applied impulses.

Consider a string fixed at both ends. Oscillations or standing waves can be established in the string. The only requirement we have to satisfy is that the end points be nodes. There may be any number of nodes in between or none at all, so that the wavelength associated with the standing waves can take on many different values. The distance between adjacent nodes is $\lambda/2$, so that in a string of length $l$ there must be exactly an integral number $n$ of half wavelengths, $\lambda/2$. That is,

$$\frac{n\lambda}{2} = l$$

or

$$\lambda = \frac{2l}{n}, \qquad n = 1, 2, 3, \cdots.$$

But $\lambda$ [illegible]$/\nu$ and $v = \sqrt{F/\mu}$, so that the natural frequencies of oscillation of the system are

$$\nu = \frac{n}{2l}\sqrt{\frac{F}{\mu}}, \qquad n = 1, 2, 3, \cdots \tag{19–19}$$

If the string is set vibrating and left to itself, the oscillations gradually die out. The motion is damped by dissipation of energy through the elastic supports at the ends and by the resistance of the air to the motion. We can pump energy into the system by applying a driving force. If the driving frequency is near that of any natural frequency of the string, the string will vibrate at that frequency with a large amplitude. Because the string has a large number of natural frequencies, resonance can occur at many different frequencies. A mass-spring system, by contrast, has only one resonant frequency. The difference is associated with the fact that in the mass-spring system the inertia characteristic is concentrated ("lumped") in one part of the system—the mass—and the elastic characteristic is concentrated in a separate part of the system—the spring. We say that this system has *lumped elements.*

A stretched string, on the other hand, is said to have *distributed elements* because every element of the string has both inertia and elastic characteristics. In the mass-spring system, there is only one way to exchange energy between kinetic and potential forms as the system oscillates; energy in kinetic form must be associated with the moving mass and energy in potential form must be associated with the deformed spring. In the stretched string, however, masslike (inertia) and springlike (elasticity) elements are distributed uniformly along the string. There are many possible ways, rather than a single way, of exchanging energy between kinetic and potential forms as the system oscillates, corresponding to the sequence of allowed values for $n$ in Eq. 19–19.

Resonance in a string is often demonstrated by attaching a string to a fixed end, by means of a weight attached to it over a pulley, and connecting the other end to a vibrator, as shown in Fig. 19–19. The transverse oscillations of the vibrator set up a traveling wave in the string which is reflected back from the fixed end. The frequency of the waves is that of the vibrator, and the wavelength is determined by $\lambda = v/\nu$. The fixed end $P$

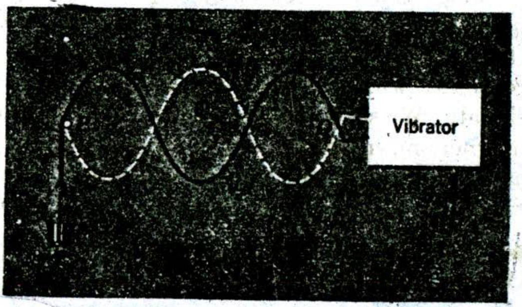


**Fig. 19–19** Standing waves in a driven string when the natural and driving frequencies are very nearly equal.

is a node, but the end $Q$ vibrates and is not. If we now vary the tension in the string by changing the hanging weight, for example, we can change the wavelength. Changing the tension changes the wave velocity, and the wavelength changes in proportion to the velocity, the frequency being constant. Whenever the wavelength becomes nearly equal to $2l/n$; where $l$ is the length of the string, we obtain standing waves of great amplitude. The string now vibrates in one of its natural modes and resonates with the vibrator. The vibrator does work on the string to maintain these oscillations against the losses due to damping. The amplitude builds up only to the point at which the vibrator expends all its energy input against damping losses. The point $Q$ is almost a node because the amplitude of the vibrator is small compared to that of the string.

Hence, with damping, the resonant frequency is almost, but not quite, a natural frequency of the string. One end point is a node, the other almost a node. In between there are points that are almost nodes, points at which the amplitude is very small. These points cannot be true nodes, for energy must flow along the string past them from the vibrator. This situation is analogous to the resonance condition for a damped harmonic oscillator with driving force, discussed in Section 15–10. There, too, the resonant frequency was almost the same as the natural frequency of the system, and the amplitude was large but not infinite. If no damping were present, the resonant frequency would be exactly a natural frequency. Then the amplitude would build up to infinity as the energy is pumped in. In practice, the system would cease to obey Hooke's law, or the small-oscillations condition, as the amplitude becomes large and the system would break. This happens even with damping, when the damping is small or the driving force is large (as in the Tacoma Bridge disaster, Fig. 15–21).

If the frequency of the vibrator is much different from a natural frequency of the system, as given by Eq. 19–19, the wave reflected at $P$ on returning to $Q$ may be much out of phase with the vibrator, and it can do work on the vibrator. That is, the string can give up some energy to the vibrator just as well as receive energy from it. The "standing" wave pattern is not fixed in form but wiggles about. On the average the amplitude is small and not much different from that of the vibrator. This situation is analogous to the erratic motion of a swing being pushed periodically with a frequency other than its natural one. The displacement of the swing is rather small.

Hence, the string absorbs peak energy from the vibrator at resonance. Tuning a radio is an analogous process. By tuning a dial the natural frequency of an alternating current in the receiving circuit is made equal to the frequency of the waves broadcast by the station desired. The circuit resonates with the transmitted signals and absorbs peak energy from the signal. We shall encounter resonance conditions again in sound, in electromagnetism, in optics, and in atomic and nuclear physics. In these areas, as in mechanics, the system will absorb peak energy from the source at resonance and relatively little energy off resonance.

▶ **Example 4.** In a demonstration with the apparatus just described, the vibrator has a frequency $\nu = 20$ cycles/sec, and the string has a linear density

$\mu = 1.56 \times 10^{-4}$ slug/ft and a length $l = 24$ ft. The tension $F$ is varied by pulling down on the end of the string over the pulley. If the demonstrator wants to show resonance, starting with one loop and then with two, three, and four loops, what force must he exert on the string?

At resonance,

$$\nu = \frac{n}{2l}\sqrt{\frac{F}{\mu}}.$$

Hence, the tension $F$ is given by

$$F = \frac{4l^2\nu^2\mu}{n^2}.$$

For one loop, $n = 1$, so that

$$F_1 = 4l^2\nu^2\mu = 4(24\text{ ft})^2(20\text{ sec}^{-1})^2(1.56 \times 10^{-4}\text{ slug/ft}) = 144\text{ lb}.$$

For two loops, $n = 2$, and

$$F_2 = \frac{4l^2\nu^2\mu}{4} = \frac{F_1}{4} = 36\text{ lb}.$$

Likewise, for three and four loops

$$F_3 = \frac{F_1}{(3)^2} = 16\text{ lb},$$

$$F_4 = \frac{F_1}{(4)^2} = 9\text{ lb}.$$

Hence, the demonstrator gradually relaxes the tension to obtain resonance with an increasing number of loops. Although the resonant frequency is always the same under these circumstances, the speed of propagation and the wavelength at resonance decrease proportionately.

Taking damping into account, are the tensions given exactly correct?

If the tension were kept fixed, giving a definite wave speed, would we obtain nore than one resonance condition by varying the frequency of the vibrator. ◀

## QUESTIONS

**1.** How could you prove experimentally that energy is associated with a wave?

**2.** Energy can be transferred by particles as well as by waves. How can we disinguish experimentally between these methods of energy transfer?

**3.** Can a wave motion be generated in which the particles of the medium vibrate with ngular simple harmonic motion? If so, explain how and describe the wave.

**4.** Are torsional waves transverse or longitudinal? Can they be considered as a uperposition of two waves, which are either transverse or longitudinal?

**5.** How can one create plane waves? Spherical waves?

**6.** The following functions in which $A$ is a constant are of the form $f(x \pm vt)$:

$$y = A(x - vt), \qquad y = A(x + vt)^2,$$
$$y = A\sqrt{x - vt}, \qquad y = A\ln(x + vt).$$

xplain why these functions are not useful in wave motion.

**7.** How do the amplitude and the intensity of surface water waves vary with the distance from the source?

**8.** The inverse square law does not apply exactly to the decrease in intensity of sounds with distance. Why not?

**9.** When two waves interfere, does one alter the progress of the other?

**10.** When waves interfere, is there a loss of energy? Explain your answer.

**11.** Why don't we observe interference effects between the light beams emitted from two flashlights or between the sound waves emitted by two violins.

**12.** If two waves differ only in amplitude and are propagated in opposite directions through a medium, will they produce standing waves? Is energy transported? Are there any nodes? (See Problem 25.)

**13.** The partial reflection of wave energy by discontinuities in the path of transmission is usually wasteful and can be minimized by insertion of "impedance matching" devices between the sections of the path bordering on the discontinuity. For example, a megaphone helps match the air column of mouth and throat to the air outside the mouth. Give other examples and explain qualitatively how such devices minimize reflection losses (see Problem 25).

**14.** Is an oscillation a wave? Explain.

**15.** Consider the standing waves in a string to be a superposition of traveling waves and explain, using superposition ideas, why there are no true nodes in the resonating string of Fig. 19–19, even at the "fixed" end. (*Hint:* Consider damping effects.)

**16.** In the discussion of transverse waves in a string we have dealt only with displacements in a single plane, the $x$-$y$ plane. If all displacements lie in one plane the wave is said to be *plane polarized.* Can there be displacements in a plane other than the single plane dealt with? If so, can two differently plane-polarized waves be combined? What appearance would such a combined wave have?

**17.** A wave transmits energy. Does it transfer momentum? Can it transfer angular momentum? (See Question 16.)

## PROBLEMS

**1.** Show that $y = y_m \sin(kx - \omega t)$ may be written in the alternative forms

$$y = y_m \sin k(x - vt), \qquad y = y_m \sin 2\pi\left(\frac{x}{\lambda} - \nu t\right),$$

$$y = y_m \sin \omega\left(\frac{x}{v} - t\right), \qquad y = y_m \sin 2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right).$$

**2.** The speed of electromagnetic waves in vacuum is $3 \times 10^8$ meters/sec. (*a*) Wavelengths in the visible part of the spectrum (light) range from about $4 \times 10^{-7}$ meter in the violet to about $7 \times 10^{-7}$ in the red. What is the range of frequencies of light waves? (*b*) The range of frequencies for shortwave radio (for example, FM radio and VHF television) is 1.5 megacycles/sec to 300 megacycles/sec. What is the corresponding wavelength range? (*c*) X-rays are also electromagnetic. Their wavelength range extends from about $5 \times 10^{-9}$ meter to $1.0 \times 10^{-11}$ meter. What is the frequency range for X-rays?

**3.** The equation of a transverse wave traveling in a rope is given by

$$y = 10 \sin \pi(0.01x - 2.00t),$$

where $y$ and $x$ are expressed in centimeters and $t$ in seconds. (*a*) Find the amplitude, frequency, velocity, and wavelength of the wave. (*b*) Find the maximum transverse speed of a particle in the rope.

**4.** Write the equation for a wave traveling in the negative direction along the $x$-axis and having an amplitude 0.010 meter, a frequency 550 vib/sec, and a speed 330 meters/sec.

**5.** A wave of frequency 500 cycles/sec has a phase velocity of 350 meters/sec. (*a*) How far apart are two points 60° out of phase? (*b*) What is the phase difference between two displacements at a certain point at times $10^{-3}$ sec apart?

**6.** (*a*) A continuous sinusoidal longitudinal wave is sent along a coil spring from a vibrating source attached to it. The frequency of the source is 25 vib/sec, and the distance between successive rarefactions in the spring is 24 cm. Find the wave speed. (*b*) If the maximum longitudinal displacement of a particle in the spring is 3.0 cm and the wave moves in the $-x$ direction, write the equation for the wave. Let the source be at $x = 0$ and the displacement at $x = 0$ and $t = 0$ be zero.

**7.** What is the speed of a transverse wave in a rope of length 2.0 meters and mass 0.060 kg under a tension of 500 nt?

**8.** Prove that the slope of a string at any point $x$ is numerically equal to the ratio of the particle speed to the wave speed at that point.

**9.** A uniform circular hoop of string is rotating clockwise in the absence of gravity (see Fig. 19–20). The tangential speed is $v_0$. Find the speed of waves traveling on this string. (*Hint:* The answer is independent of the radius of the circle and the mass per unit length of the string!)

**10.** (*a*) From Example 2 show that the *maximum* speed of a particle in a string through which a sinusoidal wave is passing is $u = y_m\omega$. (*b*) In Example 2 we saw that the particles in the string oscillate with simple harmonic motion. The mechanical energy of each particle is the sum of its potential and kinetic energies and is always equal to the *maximum* value of its kinetic energy. Consider an element of string of mass $\mu\,\Delta x$ and show that the energy *per unit length* of the string is given by

$$E_l = 2\pi^2\mu\nu^2 y_m^2.$$

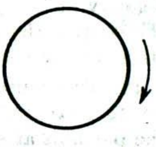


Fig. 19–20

(*c*) Show finally that the average power or average rate of transfer of energy is the product of the energy per unit length and the wave speed. (*d*) Do these results hold only for a sinusoidal wave?

**11.** Spherical waves are emitted from a 1.0-watt source in an isotropic nonabsorbing medium. What is the wave intensity 1.0 meter from the source?

**12.** (*a*) Show that the intensity $I$ (the energy crossing unit area per unit time) is the product of the energy per unit volume $e$ and the speed of propagation $v$ of a wave disturbance. (*b*) Radio waves travel at a speed of $3.0 \times 10^8$ meters/sec. Find the energy density in a radio wave 300 miles from a 50,000-watt source, assuming the waves to be spherical and the propogation to be isotropic.

**13.** A line source emits a cylindrical expanding wave. Assuming the medium absorbs no energy, find how the amplitude and intensity of the wave depend on the distance from the source.

**14.** Determine the amplitude of the resultant motion when two sinusoidal motions having the same frequency and traveling in the same direction are combined, if their amplitudes are 3.0 cm and 4.0 cm and they differ in phase by $\pi/2$ radians.

**15.** A source $S$ and a detector $D$ of high-frequency waves are a distance $d$ apart on the ground. The direct wave from $S$ is found to be in phase at $D$ with the wave from $S$

that is reflected from a horizontal layer at an altitude $H$ (Fig. 19–21). The incident and reflected rays make the same angle with the reflecting layer. When the layer rises a distance $h$, no signal is detected at $D$. Neglect absorption in the atmosphere and find the relation between $d$, $h$, $H$, and the wavelength $\lambda$ of the waves.

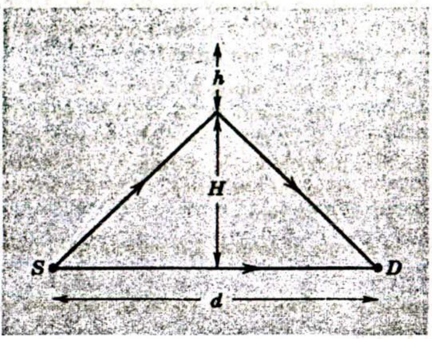


**Fig. 19–21**

**16.** Three component sinusoidal waves have the same period, but their amplitudes are in the ratio 1, 1/2, and 1/3 and their phase angles are 0, $\pi/2$, and $\pi$ respectively. Plot the resultant waveform and discuss its nature.

**17.** Four component sine waves have frequencies in the ratio 1, 2, 3, and 4 and amplitudes in the ratio 1, 1/2, 1/3, and 1/4, respectively. The first and third components are 180° out of phase with the second and fourth components. Plot the resultant waveform and discuss its nature.

**18.** A string vibrates according to the equation

$$y = 5 \sin \frac{\pi x}{3} \cos 40\pi t,$$

where $x$ and $y$ are in centimeters and $t$ is in seconds. (*a*) What are the amplitude and velocity of the component waves whose superposition can give rise to this vibration? (*b*) What is the distance between nodes? (*c*) What is the velocity of a particle of the string at the position $x = 1.5$ cm when $t = \frac{9}{8}$ sec?

**19.** Two pulses are traveling along a string in opposite direction, as shown in Fig. 19–22. (*a*) If the wave velocity is 2.0 cm/sec and the pulses are 6.0 cm apart, sketch the patterns after 0.5, 1.0, 1.5, 2.0, 2.5 sec. (*b*) What has happened to the energy at $t = 1.5$ sec?

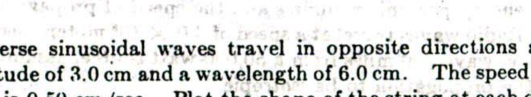


**Fig. 19–22**

**20.** Two transverse sinusoidal waves travel in opposite directions along a string. Each has an amplitude of 3.0 cm and a wavelength of 6.0 cm. The speed of a transverse wave in the string is 0.50 cm/sec. Plot the shape of the string at each of the following times: $t = 0$ (arbitrary), $t = 1.5$, $t = 3.0$, $t = 6.0$, $t = 7.5$, and $t = 9.0$ sec.

**21.** The equation of a transverse wave traveling in a rope is given by

$$y = 60 \cos \frac{\pi}{2} (0.0050x - 8.0t - 0.57),$$

in which $x$ and $y$ are expressed in centimeters and $t$ in seconds. Write down the equation of a wave that, when added to the given one, would produce standing waves on the rope.

**22.** In a laboratory experiment on standing waves a string 3.0 ft long is attached to the prong of an electrically driven tuning fork which vibrates perpendicular to the length of the string at a frequency of 60 vib/sec. The weight of the string is 0.096 lb. (*a*) What tension must the string be under (weights are attached to the other end) if it is to vibrate in four loops? (*b*) What would happen if the tuning fork is turned so as to vibrate parallel to the length of the string?

**23.** A wave travels out uniformly in all directions from a point source. Justify the following expression for the displacement $y$ of the medium at any distance $r$ from the source:

$$y = \frac{Y}{r} \sin k(r - vt).$$

Consider the speed, direction of propagation, periodicity, and intensity of the wave. What are the dimensions of the constant $Y$?

**24.** Consider two point sources $S_1$ and $S_2$ in Fig. 19–23 which emit waves of the same frequency and amplitude. The waves start in the same phase, and this phase relation at the sources is maintained throughout time. Consider points $P$ at which $r_1$ is nearly equal to $r_2$. (*a*) Show that the superposition of these two waves gives a wave whose amplitude varies with the position $P$ approximately according to

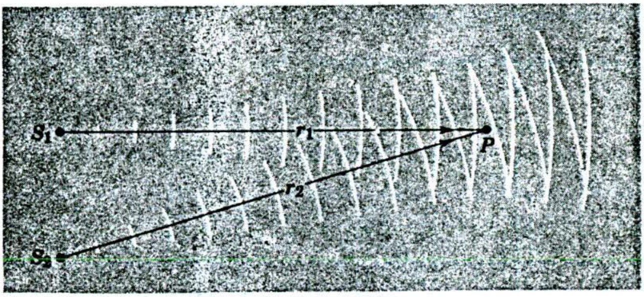


Fig. 19–23

$$\frac{2Y}{r} \cos \frac{k}{2} (r_1 - r_2),$$

in which $r = (r_1 + r_2)/2$. (*b*) Then show that total annulment occurs when $r_1 - r_2 = (n + \frac{1}{2})\lambda$, $n$ being any integer, and that total re-enforcement occurs when $r_1 - r_2 = n\lambda$.

The locus of points whose difference in distance from two fixed points is a constant is a hyperbola, the fixed points being the foci. Hence each value of $n$ gives a hyperbolic line of constructive interference and a hyperbolic line of destructive interference. At points at which $r_1$ and $r_2$ are not approximately equal (as near the sources), the amplitudes of the waves from $S_1$ and $S_2$ differ and the annulments are only partial.

**25.** If an incident traveling wave is only partially reflected from a boundary, the resulting superposition of two waves having different amplitudes and traveling in opposite directions gives a standing wave pattern of waves whose envelope is shown in Fig. 19–24. The standing wave ratio (SWR) is defined as $(A_i + A_r)/(A_i - A_r) = A_{\max}/A_{\min}$. (*a*) Show that for 100% reflection SWR $= \infty$ and that for no reflection SWR $= 1$. (*b*) Show that a measurement of the SWR just before the boundary reveals

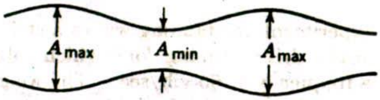


Fig. 19–24

the per cent reflection occurring at the boundary according to the formula

$$\% \text{ reflection} = [(\text{SWR} - 1)^2/(\text{SWR} + 1)^2] \times 100.$$

**26.** Two strings of linear density $\mu_1$ and $\mu_2$ are knotted together at $x = 0$ and stretched to a tension $F$. A wave $y = A \sin k_1(x - v_1t)$ in the string of density $\mu_1$ reaches the junction between the two strings, at which it is partly transmitted into the string of density $\mu_2$ and partly reflected. Call these waves $B \sin k_2(x - v_2t)$ and $C \sin k_1(x + v_1t)$, respectively. (*a*) Assuming that $k_2v_2 = k_1v_1 = \omega$ and that the displacement of the knot arising from the incident and reflected waves is the same as that arising from the transmitted wave, show that $A = B + C$. (*b*) If it is assumed that both strings near the knot have the same slope (why?), i.e., that $dy/dx$ in string 1 $= dy/dx$ in string 2, show that

$$C = A\,\frac{(k_2 - k_1)}{(k_2 + k_1)}$$

$$= A\,\frac{v_1 - v_2}{v_1 + v_2}.$$

Under what conditions is $C$ negative?

# Sound Waves

## CHAPTER 20

### 20–1 Audible, Ultrasonic, and Infrasonic Waves

Sound waves are longitudinal mechanical waves. They can be propagated in solids, liquids, and gases. The material particles transmitting such a wave oscillate in the direction of propagation of the wave itself. There is a large range of frequencies within which longitudinal mechanical waves can be generated, sound waves being confined to the frequency range which can stimulate the human ear and brain to the sensation of hearing. This range is from about 20 cycles/sec to about 20,000 cycles/sec and is called the *audible* range. A longitudinal mechanical wave whose frequency is below the audible range is called an *infrasonic* wave, and one whose frequency is above the audible range is called an *ultrasonic* wave.

Infrasonic waves of interest are usually generated by large sources, earthquake waves being an example.* The high frequencies associated with ultrasonic waves may be produced by elastic vibrations of a quartz crystal induced by resonance with an applied alternating electric field (piezoelectric effect). It is possible to produce ultrasonic frequencies as high as $6 \times 10^8$ cycles/sec in this way; the corresponding wavelength in air is about $5 \times 10^{-5}$ cm, the same as the length of visible light waves.

Audible waves originate in vibrating strings (violin, human vocal cords), vibrating air columns (organ, clarinet), and vibrating plates and membranes (xylophone, loudspeaker, drum). All of these vibrating elements alternately compress the surrounding air on a forward movement and rarefy it on a backward movement. The air transmits these disturbances

---

* See "Long Earthquake Waves," by Jack Oliver, *Scientific American*, March 1959.

outward from the source as a wave. Upon entering the ear, these waves produce the sensation of sound. Waveforms which are approximately periodic or consist of a small number of approximately periodic components give rise to a pleasant sensation (if the intensity is not too high), as, for example, musical sounds. Sound whose waveform is nonperiodic is heard as noise. Noise can be represented as a superposition of periodic waves, but the number of components is very large.

In this chapter we deal with the properties of longitudinal mechanical waves, using sound waves as the prototype.

## 20–2 Propagation and Speed of Longitudinal Waves

Sound waves, if unimpeded, will spread out in all directions from a source. It is simpler to deal with one-dimensional propagation, however, than with three-dimensional propagation, so that we consider first the transmission of longitudinal waves in a tube.

Figure 20–1 shows a piston at one end of a long tube filled with a compressible medium. The vertical lines divide the compressional (fluid) medium into thin "slices," each of which contains the same mass of fluid. Where the lines are relatively close together the fluid pressure and density are greater than they are in the normal undisturbed fluid, and conversely. We shall treat the fluid as a continuous medium and ignore for the time being the fact that it is made up of molecules that are in continual random motion.

If we push the piston of Fig. 20–1 forward, the fluid in front of it is compressed, the fluid pressure and density rising above their normal undisturbed values. The compressed fluid moves forward, compressing the fluid layers next to it, and a compressional pulse travels down the tube. If we then withdraw the piston, the fluid in front of it expands, its pressure and density falling below their normal undisturbed values; a pulse of rarefaction travels down the tube. These pulses are similar to transverse pulses traveling along a string, except that the oscillating fluid elements are

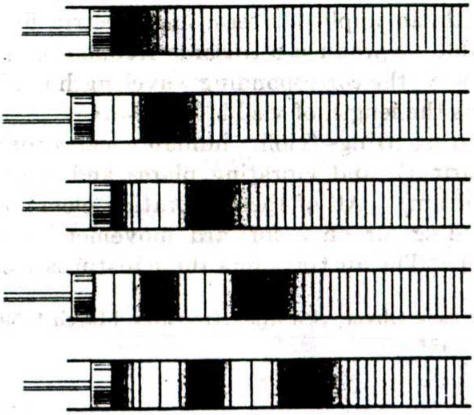


**Fig. 20–1** Sound waves generated in a tube by an oscillating piston. The vertical lines divide the compressible medium in the tube into layers of equal mass.

displaced along the direction of propagation (longitudinal) instead of at right angles to this direction (transverse). If the piston oscillates back and forth, a continuous train of compressions and rarefactions will travel along the tube (Fig. 20–1). As for transverse waves in a string (see Section 19–5) we should be able, using Newton's laws of motion, to express the speed of propagation of this longitudinal wave in terms of an elastic and an inertial property of the medium. We now do so.

For the moment, let us assume that the tube is very long so that we can ignore reflections from the other end. As for the string of Fig. 19–6, we will consider not an extended wave but a single (compressional) pulse that we might generate by giving the piston in Fig. 20–1 a short, rapid, inward stroke.

Figure 20–2 shows such a pulse (labeled "compressional zone") traveling at speed $v$ along the tube from left to right. For simplicity we have assumed this pulse to have sharply defined leading and trailing edges and to have a uniform fluid pressure and density in its interior. When we analyzed the motion of a transverse pulse in a string, we found it convenient to choose a reference frame in which the pulse remained stationary; we will do this here also. In Fig. 20–2, then, the compressional zone remains stationary in our reference frame while the fluid moves through it from right to left with speed $v$, as shown.

Let us follow the motion of the element of fluid contained between the vertical lines at $P$ in Fig. 20–2. This element moves forward at speed $v$ until it strikes the compressional zone. While it is entering this zone it encounters a difference of pressure $\Delta p$ between its leading and its trailing edges. The element is compressed and *decelerated*, moving with a lower speed $v + \Delta v$ within the zone, the quantity $\Delta v$ being negative. The element eventually emerges from the left face of the zone where it expands to its original volume and the pressure differential $\Delta p$ acts to *accelerate* it to its original speed $v$. The figure shows the element at point $R$, having passed through the compressional zone and moving again with speed $v$, as at $P$.

Let us apply Newton's laws to the fluid element while it is entering the compressional zone. The resultant force acting during entry points to the right in Fig. 20–2 and has magnitude

$$F = (p + \Delta p)A - pA = \Delta pA$$

in which $A$ is the cross-sectional area of the tube.

The length of the element outside the compressional zone (at $P$, say) is $v\,\Delta t$, where $\Delta t$ is the time required for the element to move past any given point. The volume of the element is thus $vA\,\Delta t$ and its mass is $\rho_0 vA\,\Delta t$, where $\rho_0$ is the density of the fluid outside the compressional zone. The deceleration $a$ experienced by the

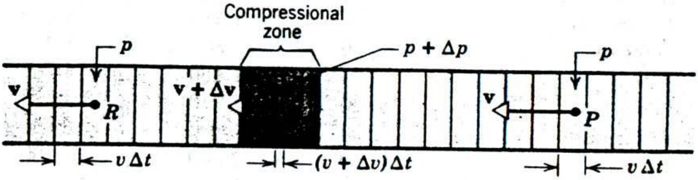


**Fig. 20–2** A compressional pulse travels along a gas-filled tube. In a reference frame in which the undisturbed gas is at rest the pulse moves from left to right with speed $v$. We view the pulse, however, from a reference frame in which the pulse is stationary; in such a frame the gas outside the pulse streams through the tube from right to left with speed $v$, as shown. Note that $\Delta v$ is negative.

element as it enters the zone is $-\Delta v/\Delta t$; because $\Delta v$ is inherently negative, $a$ is positive, which means that, like the force $\Delta pA$ in Fig. 20–2, it points to the right. Thus Newton's second law

$$F = ma$$

yields

$$\Delta pA = (\rho_0 vA\,\Delta t)\,\frac{-\Delta v}{\Delta t},$$

which we may write as

$$\rho_0 v^2 = \frac{-\Delta p}{\Delta v/v}.$$

Now the fluid that would occupy a volume $V = Av\,\Delta t$ at $P$ is compressed by an amount $A(\Delta v)\,\Delta t = \Delta V$ on entering the compressional zone. Hence,

$$\frac{\Delta V}{V} = \frac{A\,\Delta v\,\Delta t}{Av\,\Delta t} = \frac{\Delta v}{v}$$

and we obtain

$$\rho_0 v^2 = \frac{-\Delta p}{\Delta V/V}.$$

The ratio of the change in pressure on a body, $\Delta p$, to the fractional change in volume resulting, $-\Delta V/V$, is called the *bulk modulus of elasticity* $B$ of the body. That is, $B = -V\,\Delta p/\Delta V$. $B$ is positive because an increase in pressure causes a decrease in volume. In terms of $B$, the speed of the longitudinal pulse in the medium of Fig. 20–2 is

$$v = \sqrt{B/\rho_0}. \tag{20–1}$$

A more extended analysis than given above shows that Eq. 20–1 applies not only to rectangular pulses of the type displayed in Fig. 20–2 but also

**Table 20–1**

SPEED OF SOUND

| Medium | Temperature, °C | Speed | |
|---|---|---|---|
| | | Meters/sec | Ft/sec |
| Air | 0 | 331.3 | 1,087 |
| Hydrogen | 0 | 1,286 | 4,220 |
| Oxygen | 0 | 317.2 | 1,041 |
| Water | 15 | 1,450 | 4,760 |
| Lead | 20 | 1,230 | 4,030 |
| Aluminum | 20 | 5,100 | 16,700 |
| Copper | 20 | 3,560 | 11,700 |
| Iron | 20 | 5,130 | 16,800 |
| Extreme values | | | |
| Granite | | 6,000 | 19,700 |
| Vulcanized rubber | 0 | 54 | 177 |

to pulses of any shape and to extended wave trains. Notice that the speed of the wave is determined by the properties of the medium through which it propagates, and that an elastic property $B$ and an inertial property $\rho_0$ are involved. Table 20-1 gives the speed of longitudinal (sound) waves in various media.

If the medium is a gas, such as air, it is possible to express $B$ in terms of the undisturbed gas pressure $p_0$. For a sound wave in a gas we obtain

$$v = \sqrt{\gamma p_0/\rho_0},$$

where $\gamma$ is a constant called the ratio of specific heats for the gas (Chapter 23).

If the medium is a solid, for a thin rod the bulk modulus is replaced by a stretch modulus (called Young's modulus). If the solid is extended, we must allow for the fact that, unlike a fluid, a solid offers elastic resistance to tangential or shearing forces and the speed of longitudinal waves will depend on the shear modulus as well as the bulk modulus.

## 20-3 Traveling Longitudinal Waves

Consider again the continuous train of compressions and rarefactions traveling down the tube of Fig. 20-1. As the wave advances along the tube, each small volume element of fluid oscillates about its equilibrium position. The displacement is to the right or left along the $x$-direction of propagation of the wave. For convenience let us represent the displacement of any such volume element (or layer of elements that move in the same way) from its equilibrium position at $x$ by the letter $y$. It is to be understood that the displacement $y$ is *along the direction of propagation* for a longitudinal wave, whereas for a transverse wave the displacement $y$ is *at right angles to the direction of propagation.* Then the equation of a longitudinal wave traveling to the right may be written as

$$y = f(x - vt).$$

For the particular case of a simple harmonic oscillation we may have

$$y = y_m \cos \frac{2\pi}{\lambda} (x - vt).$$

In this equation $v$ is the speed of the longitudinal wave, $y_m$ is its amplitude, and $\lambda$ is its wavelength; $y$ gives the displacement of a particle at time $t$ from its equilibrium position at $x$. As before, we may write this more compactly as

$$y = y_m \cos (kx - \omega t). \tag{20-2}$$

It is usually more convenient to deal with pressure variations in a sound wave than with the actual displacements of the particles conveying the wave. Let us therefore write the equation of the wave in terms of the pressure variation rather than in terms of the displacement.

From the relation

$$B = -\frac{\Delta p}{\Delta V/V},$$

we have

$$\Delta p = -B\frac{\Delta V}{V}.$$

Just as we let $y$ represent the displacement from the equilibrium position $x$, so we now let $p$ represent the *change* from the undisturbed pressure $p_0$. Then $p$ replaces $\Delta p$, and

$$p = -B\frac{\Delta V}{V}.$$

If a layer of fluid at pressure $p_0$ has a thickness $\Delta x$ and cross-sectional area $A$, its volume is $V = A\,\Delta x$. When the pressure changes, its volume will change by $A\,\Delta y$, where $\Delta y$ is the amount by which the thickness of the layer changes during compression or rarefaction. Hence,

$$p = -B\frac{\Delta V}{V} = -B\frac{A\,\Delta y}{A\,\Delta x}.$$

As we let $\Delta x \to 0$ so as to shrink the fluid layer to infinitesimal thickness, we obtain

$$p = -B\frac{\partial y}{\partial x}. \qquad (20\text{–}3)$$

We have used partial derivative notation because (see Eq. 20–2) $y$ is a function of both $x$ and $t$ and we take the latter quantity as constant in this discussion. If the particle displacement is simple harmonic, then, from Eq. 20–2, we obtain

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

and from Eq. 20–3

$$p = Bky_m \sin(kx - \omega t). \qquad (20\text{–}4)$$

Hence, the pressure variation at each position $x$ is also simple harmonic.

Since $v = \sqrt{B/\rho_0}$, we can write Eq. 20–4 more conveniently as

$$p = [k\rho_0 v^2 y_m]\sin(kx - \omega t).$$

Recall that $p$ represents the change from standard pressure $p_0$. The term in brackets represents the maximum change in pressure and is called the *pressure amplitude*. If we denote this by $P$, then

$$p = P\sin(kx - \omega t), \qquad (20\text{–}5)$$

where

$$P = k\rho_0 v^2 y_m. \qquad (20\text{–}6)$$

Hence, a sound wave may be considered either as a displacement wave or as a pressure wave. If the former is written as a cosine function, the latter will be a sine function and vice versa. The displacement wave is thus 90° out of phase with the pressure wave. That is, when the displacement from equilibrium at a point is a maximum or a minimum, the excess pressure there is zero; when the displacement at a point is zero, the excess or deficiency of pressure there is a maximum. Equation 20–6 gives the relation between the pressure amplitude (maximum variation of pressure from equilibrium) and the displacement amplitude (maximum variation of position from equilibrium). The student should check the dimensions of each side of Eq. 20–6 for consistency. What units may the pressure amplitude have?

The intensity of a wave is proportional to the square of the displacement amplitude of the wave; see Section 19–6. We have just shown that for sound waves the pressure amplitude is proportional to the displacement amplitude. Hence, the intensity of a sound wave is proportional to the square of the pressure amplitude. In fact, when the intensity is expressed in terms of the pressure amplitude, the frequency does not appear explicitly in the expression (see Problem 9). Hence, by measuring pressure changes, the intensities of sounds having *different* frequencies can be compared directly. For this reason instruments that measure pressure changes are preferred to those that measure displacement amplitude. As we shall see in Example 1, the displacement amplitudes would be difficult to measure in any case.

▶ **Example 1.** (*a*) The maximum pressure variation $P$ that the ear can tolerate in loud sounds is about 28 nt/meter$^2$. Normal atmospheric pressure is about 100,000 nt/meter$^2$. Find the corresponding maximum displacement for a sound wave in air having a frequency of 1000 cycles/sec.

From Eq. 20–6 we have

$$y_m = \frac{P}{k\rho_0 v^2}.$$

From Table 20–1, $v = 331$ meters/sec so that

$$k = \frac{2\pi}{\lambda} = \frac{2\pi\nu}{v} = \frac{2\pi \times 10^3}{331}\ \text{meter}^{-1} = 19\ \text{meter}^{-1}.$$

The density of air $\rho_0$ is 1.22 kg/meter$^3$. Hence, for $P = 28$ nt/meter$^2$ we obtain

$$y_m = \frac{28}{(19)(1.22)(331)^2}\ \text{meter} = 1.1 \times 10^{-5}\ \text{meter}.$$

The displacement amplitudes for the *loudest* sounds are about $10^{-5}$ meter, a very small value indeed.

(*b*) In the faintest sound that can be heard at 1000 cycles/sec the pressure amplitude is about $2.0 \times 10^{-5}$ nt/meter$^2$. Find the corresponding displacement amplitude.

From $y_m = P/k\rho_0 v^2$, using these values for $k$, $v$, and $\rho_0$, we obtain, with $P = 2.0 \times 10^{-5}$ nt/meter$^2$,

$$y_m \cong 8 \times 10^{-12}\ \text{meter} \cong 10^{-11}\ \text{meter}.$$

This is smaller than the radius of an atom, which is about $10^{-10}$ meter! How can it be that the ear responds to such a small displacement? ◀

In our analysis we have ignored the molecular structure of matter and treated the fluid as a continuous medium. In gases, however, the spaces between molecules are large compared to the diameters of the molecules. The molecules move about at random. The oscillations produced by a sound wave passing through are superimposed on this random thermal motion. An impulse given to one molecule is passed on to another molecule only after the first one has moved through the empty space between them and collided with the second. From this brief discussion, would you ever expect the speed of sound to exceed the average molecular speed in a fluid?

## 20–4 Standing Longitudinal Waves

Longitudinal waves traveling along a tube are reflected at the ends of the tube, just as transverse waves in a string are reflected at its ends. Interference between the waves traveling in opposite directions gives rise to standing longitudinal waves.

If the end of the tube is closed, the reflected wave is 180° out of phase with the incident wave. This result is a necessary consequence of the fact that the displacement of the small volume elements at a closed end must always be zero. Hence, a closed end is a displacement *node.* If the end of the tube is open, the fluid elements there are free to move. However, the nature of the reflection there depends on whether the tube is wide or narrow compared to the wavelength. If the tube is narrow compared to the wavelength, as in most musical instruments, the reflected wave has nearly the same phase as the incident wave. Then the open end is almost a displacement *antinode.* The exact antinode is usually somewhere near the opening, but the effective length of the air columns of a wind instrument, for example, is not as definite as the length of a string fixed at both ends.

Standing longitudinal waves in a gas column can be dramatically demonstrated by means of the apparatus shown in Fig. 20–3. A source of longitudinal waves, such as the speaker of an audio oscillator at $S$, sets up vibrations in a flexible diaphragm at one end of the tube. Gas fills the tube from the inlet and passes slowly out through regularly spaced small openings along the top. The escaping gas is lit, giving a series of flames. When a frequency is found at which the gas column is in resonance, the amplitude of the standing longitudinal waves becomes rather large and we can see a wavelike variation in the height and width of the gas flames along the tube. The interval between nodes or antinodes is clearly visible. By varying

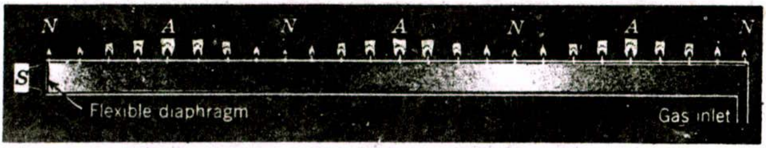


**Fig. 20–3** Flames show the presence of standing waves in a tube filled with illuminating gas. $A$ and $N$ refer to *displacement* antinodes and nodes, respectively.

the frequency we can pass from one resonance condition to another. The natural modes of oscillation of the gas column are determined by the effective length of the column and the wave speed. The wavelength $\lambda$ at resonance can be taken to be twice the distance between adjacent nodes (or antinodes), and knowing the frequency $\nu$ of the source at resonance, we can determine the wave speed in the gas under these conditions from $v = \nu\lambda$. In practice there are more flexible and accurate ways to measure the speed of sound in gases. (See Problem 17 and Example 2.)

In Fig. 20–3 the nodes and antinodes, $N$ and $A$, refer to the particle *displacements* in the standing wave. At a displacement node, the pressure variations (above and below the average) are a maximum. Hence, a displacement node corresponds to a pressure antinode. At a displacement antinode the pressure remains constant with time. Hence, a displacement antinode corresponds to a pressure node.

This can be understood physically by realizing that two small volume elements of gas on opposite sides of a displacement node are vibrating in *opposite phase*. Hence, when they approach each other, the pressure at this node is a maximum, and when they recede from each other, the pressure at this node is a minimum. Two small elements of gas which are on opposite sides of a displacement antinode vibrate *in phase* and therefore give rise to no pressure variations at the antinode.

## 20–5 Vibrating Systems and Sources of Sound

If a string fixed at both ends is bowed, transverse vibrations travel along the string; these disturbances are reflected at the fixed ends, and a standing wave pattern is formed. The natural modes of vibration of the string are excited and these vibrations give rise to longitudinal waves in the surrounding air which transmits them to our ears as a musical sound.

We have seen (Section 19–10) that a string of length $l$, fixed at both ends, can resonate at frequencies given by

$$\nu_n = \frac{n}{2l}\,v = \frac{n}{2l}\sqrt{\frac{F}{\mu}}, \qquad n = 1, 2, 3, \ldots . \tag{20–7}$$

Here $v$ is the speed of the transverse waves in the string whose superposition can be thought of as giving rise to the vibrations; the speed $v$ $(= \sqrt{F/\mu})$ is the same for all frequencies. At any one of these frequencies the string will contain a whole number $n$ of loops between its ends, and the condition that the ends be nodes is met (Fig. 20–4).

The lowest frequency, $\sqrt{F/\mu}/2l$, is called the *fundamental* frequency $\nu_1$ and the others are called *overtones*. Overtones whose frequencies are integral multiples of the fundamental are said to form a harmonic series. The fundamental is the first harmonic. The frequency $2\nu_1$ is the first overtone or the second harmonic, the frequency $3\nu_1$ is the second overtone or the third harmonic, and so on.

If the string is initially distorted so that its shape is the same as ***any one***

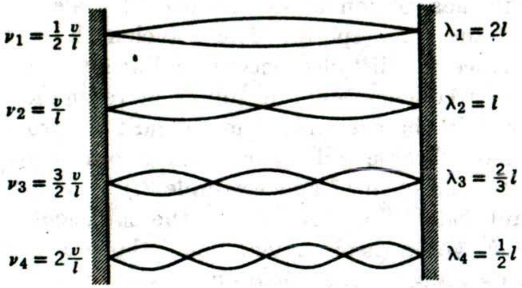


**Fig. 20–4** The first four modes of vibration of a string fixed at both ends. Note that $\nu_n\lambda_n = v = \sqrt{F/\mu}$.

of the possible harmonics, it will vibrate at the frequency of that particular harmonic, when released. The initial conditions usually arise from striking or bowing the string, however, and in such cases not only the fundamental but many of the overtones are present in the resulting vibration. We have a superposition of several natural modes of oscillation. The actual displacement is the sum of the several harmonics with various amplitudes; see Fig. 19–12. The impulses that are sent through the air to the ear and brain give rise to one net effect which is characteristic of the particular stringed instrument. The quality of the sound of a particular note (fundamental frequency) played by an instrument is determined by the number of overtones present and their respective intensities. Figure 20–5 shows the sound spectra and corresponding waveforms for the violin and piano.*

An organ pipe is a simple example of sound originating in a vibrating air column. If both ends of a pipe are open and a stream of air is directed against an edge, standing longitudinal waves can be set up in the tube.

---

* See "The Physics of the Piano" by E. Donnell Blackham in *Scientific American*, December 1965.

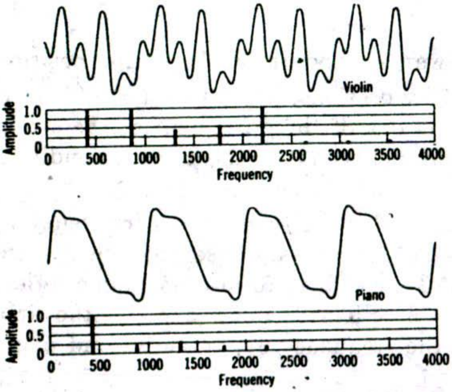


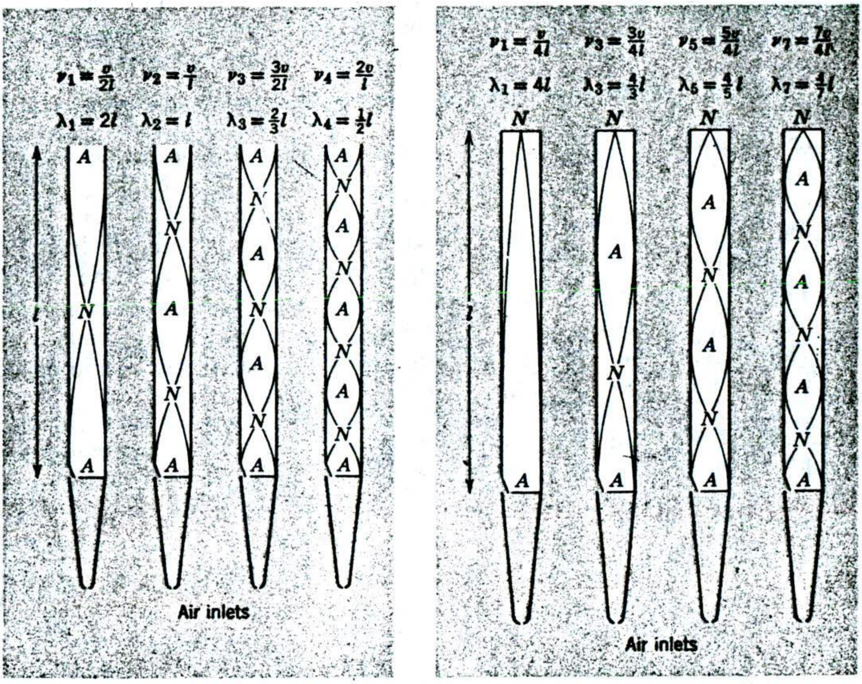
**Fig. 20–5** Waveform and sound spectrum for two stringed instruments, the violin and the piano. The fundamental frequency in both cases is 440 cycles/sec (concert A). In each diagram we show only four cycles of the wave. The sound spectrum shows the relative amplitude of the various harmonic components of the wave. Notice the presence of loud higher harmonics (especially the fifth) in the violin spectrum.

The air column will then resonate at its natural frequencies of vibration, given by

$$\nu_n = \frac{n}{2l}\,v, \qquad n = 1, 2, 3, \ldots .$$

Here $v$ is the speed of the longitudinal waves in the column whose superposition can be thought of as giving rise to the vibrations, and $n$ is the number of half wavelengths in the length $l$ of the column. As with the bowed string, the fundamental and overtones are excited at the same time.

In an open pipe the fundamental frequency corresponds (approximately) to a displacement antinode at each end and a displacement node in the middle, as shown in Fig. 20–6*a*. The succeeding drawings of Fig. 20–6*a* show three of the overtones, the second, third, and fourth harmonics. Hence, in an open pipe the fundamental frequency is $v/2l$ and *all* harmonics are present.



(*a*) (*b*)

**Fig. 20–6** (*a*) The first four modes of an open organ pipe. The distance from the center line of the pipe to the light lines drawn inside the pipe shows the displacement amplitude at each place. $N$ and $A$ mark the locations of the displacement nodes and antinodes. Note that *both* ends of the pipe are open. (*b*) The first four modes of vibration of a closed organ pipe. Notice that the even-numbered harmonics are absent and the upper end of the pipe is closed.

In a closed pipe the closed end is a displacement node. Figure 20–6*b* shows the modes of vibration of a closed pipe. The fundamental frequency is $v/4l$ (approximately), which is one-half that of an open pipe of the same length. The only overtones present are those that give a displacement node at the closed end and an antinode (approximately) at the open end. Hence, as is shown in Fig. 20–6*b*, the second, fourth, etc., harmonics are missing. In a closed pipe the fundamental frequency is $v/4l$, and only the *odd* harmonics are present. The quality of the sounds from an open pipe is therefore different from that from a closed pipe.

Vibrating rods, plates, and stretched membranes also give rise to sound waves. Consider a stretched flexible membrane, such as a drumhead. If it is struck a blow, a two-dimensional pulse travels outward from the struck point and is reflected again and again at the boundary of the membrane. If some point of the membrane is forced to vibrate periodically, continuous trains of waves travel out along the membrane. Just as in the one-dimensional case of the string, so here too standing waves can be set up in the two-dimensional membrane. Each of these standing waves has a certain frequency natural to (or characteristic of) the membrane. Again the lowest frequency is called the fundamental and the others are overtones. Generally, a number of overtones are present along with the fundamental when the membrane is vibrating. These vibrations may excite sound waves of the same frequency.

The nodes of a vibrating membrane are lines rather than points (as in a vibrating string) or planes (as in a pipe). Since the boundary of the membrane is fixed, it must be a nodal line. For a circular membrane fixed at its edge, possible modes of vibration along with their nodal lines are shown in Fig. 20–7. The natural frequency of each mode is given in terms of the fundamental $\nu_1$. Notice that the frequencies of the overtones are *not* harmonics, that is, they are not integral multiples of $\nu_1$. Vibrating rods also have a nonharmonic set of natural frequencies. Rods and plates have limited use as musical instruments for this reason.

In general, we find that all elastic bodies will vibrate freely with a definite set of frequencies for a given set of boundary or end conditions. These frequencies are called proper frequencies, characteristic frequencies, or *eigenfrequencies** of the system. In general, the eigenfrequencies do *not* form a harmonic series, although some of them may be related as the ratio of whole numbers. In all these cases we have standing waves, and certain regions of the bodies stay at rest all the time. These nodes are curves in two-dimensional bodies and surfaces in three-dimensional bodies.

Recall that for a vibrating string the equation describing a standing wave (see Eq. 19-18*b*) is of the type

$$y = 2y_m \cos 2\pi\nu t \sin \frac{2\pi x}{\lambda}.$$

This holds for a string fixed at both ends ($y = 0$ at $x = 0$ and $x = n\lambda/2$). The

---

* *Eigen*—from the German—meaning *own, individual, characteristic.*

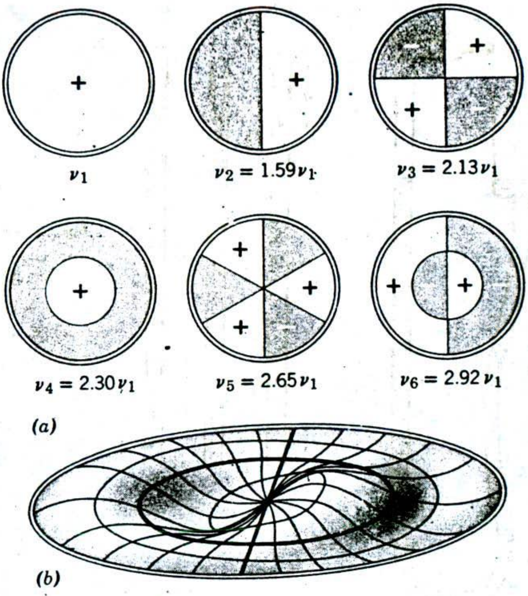


**Fig. 20–7** (*a*) The first six modes of vibration of a circular drumhead clamped around its periphery. The lines represent nodes, the circumference being a node in every case. The + and − signs represent opposite displacements; at an instant when the + areas are raised, the − areas will be depressed. Note that the frequency of each mode is not an integral multiple of the fundamental $\nu_1$ as is the case for strings and tubes. (*b*) A sketch of a drum-head vibrating in mode $\nu_6$. The displacement shown here is exaggerated for clarity.

picture of the string at any time is determined by the equation

$$y = C \sin \frac{2\pi x}{\lambda} = C \sin \frac{n\pi x}{l} \qquad (t = \text{constant}),$$

where $C$ is a constant "scale factor," whose value varies with time; $l$ is the length of the string, and $n$ is an integer specifying the mode of vibration (the harmonic). This function $\sin 2\pi x/\lambda$ fixes the position of the nodes and is called the proper function, characteristic function, or *eigenfunction* of the string.

Likewise, the nodes of *any* vibrating elastic body are fixed by certain functions of position which are called the eigenfunctions of the problem. In general, these functions are *not* sinusoidal functions but are functions that become zero for certain values of the coordinates. The determination of these functions and the corresponding values of the eigenfrequencies is an important problem in atomic, nuclear, and solid-state physics. They characterize the behavior of such systems. It is in quantum mechanics that the procedure has been successfully worked out for microscopic systems. However, the results bear a striking analogy to the results of classical vibration and wave theory, as applied to macroscopic systems.

▶ **Example 2.** Figure 20–8 shows a simple apparatus that can be used to measure the speed of sound in air by resonance methods. A vibrating tuning fork of frequency $\nu$ is held near the open end of a tube. The tube is partly filled with water. The length of the air column can be varied by changing the water level. It is found that the sound intensity is a maximum when the water level is gradually lowered from the top of the tube a distance $a$. Thereafter, the intensity reaches a maximum again at distances $s$, $2s$, $3s$, etc., below the level at $a$. Find the speed of sound in air.

The sound intensity reaches a maximum when the air column resonates with the tuning fork. The air column acts like a tube closed at one end. The standing wave pattern consists of a node at the water surface and an antinode near the

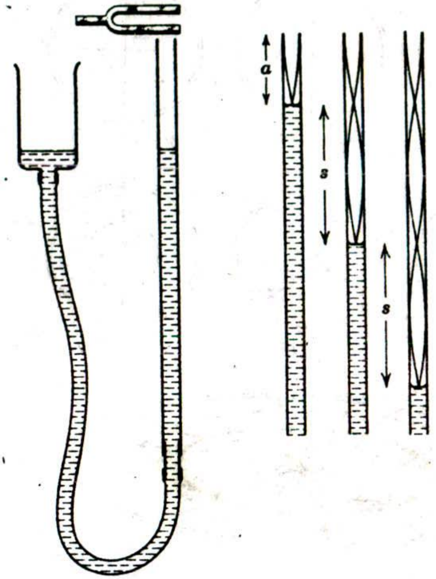


**Fig. 20–8** Example 2. Measuring the speed of sound in air. The water level in the tube can be adjusted by raising or lowering the reservoir on the left which is connected to the tube by a rubber hose.

open end. Since the frequency of the source is fixed and the speed of sound in the air column has a definite value, resonance occurs at one specific wavelength,

$$\lambda = \frac{v}{\nu}.$$

The distance $s$ between successive resonance positions is therefore the distance between adjacent nodes. (See Fig. 20–8.) Hence,

$$s = \frac{\lambda}{2} \qquad \text{or} \qquad \lambda = 2s.$$

Combining equations we find

$$2s = \frac{v}{\nu} \qquad \text{or} \qquad v = 2s\nu.$$

In an experiment with a fork of frequency $\nu = 1080$ cycles/sec, $s$ is found to be 15.3 cm. Hence,

$$\lambda = 2s = 30.6 \text{ cm}$$

and $$v = \nu\lambda = (1080)(0.306) \text{ meters/sec} = 330 \text{ meters/sec}.$$

What significance does the distance $a$ have? Could gases other than air be used conveniently in this apparatus? ◀

## 20–6 Beats

When two wavetrains of the same frequency travel along the same line in opposite directions, standing waves are formed in accord with the principle of superposition. We may characterize these waves by drawing a plot of the amplitude of oscillation as a function of distance, as in Fig. 20-4. This illustrates a type of interference that we can call *interference in space.*

The same principle of superposition leads us to another type of interference, which we can call *interference in time.* It occurs when two wavetrains of slightly *different frequency* travel through the same region. With sound such a condition exists when, for example, two adjacent piano keys are struck simultaneously.

Consider some one point in space through which the waves are passing. In Fig. 20-9*a* we plot the displacements produced at such a point by the two waves separately as a function of time. For simplicity we have assumed that the two waves have equal amplitude, although this is not necessary. The resultant vibration at that point as a function of time is the sum of the individual vibrations and is plotted in Fig. 20-9*b*. We see that the *amplitude* of the resultant wave at the given point is not constant but *varies in time.* In the case of sound the varying amplitude gives rise to variations in loudness which are called *beats.* Two strings may be tuned to the same frequency by tightening one of them while sounding both until the beats disappear.

Let us represent the displacement at the point produced by one wave as

$$y_1 = y_m \cos 2\pi\nu_1 t,$$

and the displacement at the point produced by the other wave of equal amplitude as

$$y_2 = y_m \cos 2\pi\nu_2 t.$$

By the superposition principle, the resultant displacement is

$$y = y_1 + y_2 = y_m(\cos 2\pi\nu_1 t + \cos 2\pi\nu_2 t),$$

and since

$$\cos a + \cos b = 2 \cos \frac{a - b}{2} \cos \frac{a + b}{2},$$

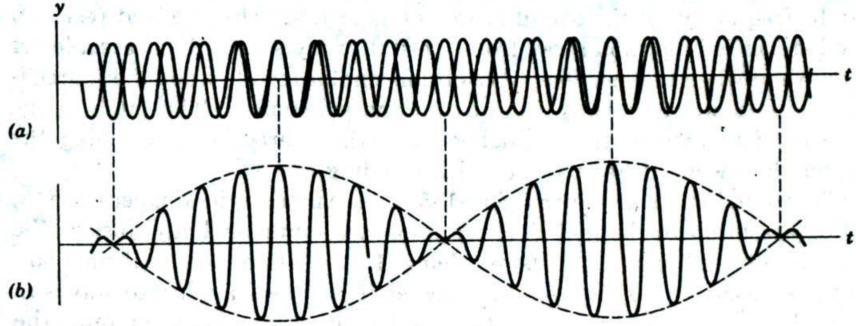


**Fig. 20-9** The beat phenomenon. Two waves of slightly different frequencies, shown in (*a*), combine in (*b*) to give a wave whose amplitude (dashed line) varies periodically with time. Compare with Fig. 19-14, which shows the same phenomenon displayed as a function of dist

this can be written as

$$y = \left[2y_m \cos 2\pi \left(\frac{\nu_1 - \nu_2}{2}\right) t\right] \cos 2\pi \left(\frac{\nu_1 + \nu_2}{2}\right) t. \qquad (20\text{–}8)$$

The resulting vibration may then be considered to have a frequency

$$\bar{\nu} = \frac{\nu_1 + \nu_2}{2},$$

which is the average frequency of the two waves, and an amplitude given by the expression in brackets. Hence, the amplitude itself varies with time with a frequency

$$\nu_{\text{amp}} = \frac{\nu_1 - \nu_2}{2}.$$

If $\nu_1$ and $\nu_2$ are nearly equal, this term is small and the amplitude fluctuates slowly. This phenomenon is a form of amplitude modulation which has a counterpart (side bands) in AM radio receivers.

A beat, that is, a maximum of amplitude, will occur whenever

$$\cos 2\pi \left(\frac{\nu_1 - \nu_2}{2}\right) t$$

equals 1 or $-1$. Since *each* of these values occurs once in each cycle (see Fig. 19–14), the number of beats per second is *twice* the frequency $\nu_{\text{amp}}$ or $\nu_1 - \nu_2$. Hence, the number of beats per second equals the difference of the frequencies of the component waves. Beats between two tones can be detected by the ear up to a frequency of about seven per second. At higher frequencies individual beats cannot be distinguished in the sound produced.

## 20–7 The Doppler Effect

When a listener is in motion toward a stationary source of sound, the pitch (frequency) of the sound heard is higher than when he is at rest. If the listener is in motion away from the stationary source, he hears a lower pitch than when he is at rest. We obtain similar results when the source is in motion toward or away from a stationary listener. The pitch of the whistle of the locomotive is higher when the source is approaching the hearer than when it has passed and is receding.

Christian Johann Doppler (1803–1853), an Austrian, in a paper of 1842, called attention to the fact that the color of a luminous body, just as the pitch of a sounding body, must be changed by relative motion of the body and the observer. This *Doppler effect*, as it is called, applies to waves in general. Let us apply it now to sound waves. We consider only the special case in which the source and observer move along the line joining them.

Let us consider a reference frame at rest in the medium through which the sound travels. Figure 20–10 shows a source of sound $S$ at rest in this

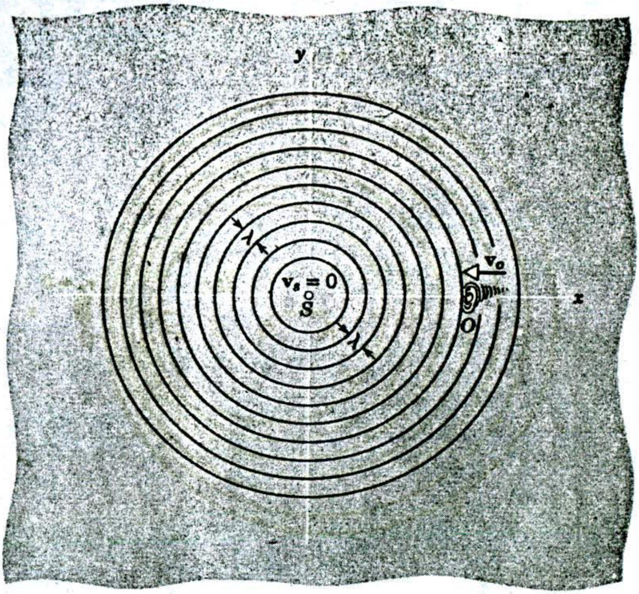


**Fig. 20–10** The Doppler effect due to motion of the observer (ear). The source is at rest.

frame and an observer $O$ (note the ear) moving *toward* the source at a speed $v_o$. The circles represent wavefronts, spaced one wavelength apart, traveling through the medium. If the observer were at rest in the medium he would receive $vt/\lambda$ waves in time $t$, where $v$ is the speed of sound in the medium and $\lambda$ is the wavelength. Because of his motion toward the source, however, he receives $v_ot/\lambda$ *additional* waves in this same time $t$. The frequency $\nu'$ that he hears is the number of waves received per unit time or

$$\nu' = \frac{vt/\lambda + v_ot/\lambda}{t} = \frac{v + v_o}{\lambda} = \frac{v + v_o}{v/\nu}.$$

That is,

$$\nu' = \nu\frac{v + v_o}{v} = \nu\left(1 + \frac{v_o}{v}\right). \tag{20–9a}$$

The frequency $\nu'$ heard by the observer is the ordinary frequency $\nu$ heard at rest plus the increase $\nu(v_o/v)$ arising from the motion of the observer. When the observer is in motion *away* from the stationary source, there is a *decrease* in frequency $\nu(v_o/v)$ corresponding to the waves that do not reach the observer each unit of time because of his receding motion. Then

$$\nu' = \nu\left(\frac{v - v_o}{v}\right) = \nu\left(1 - \frac{v_o}{v}\right). \tag{20–9b}$$

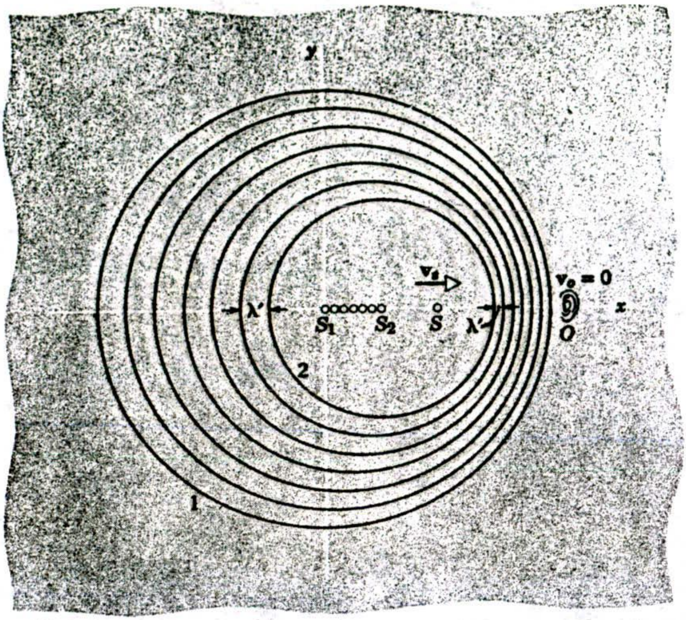


**Fig. 20–11** The Doppler effect due to motion of the source. The observer is at rest. Wavefront 1 was emitted by the source when it was at $S_1$, wavefront 2 was emitted when it was at $S_2$, etc. At the instant the "snapshot" was taken, the source was at $S$.

Hence, the general relation holding when the *source is at rest* with respect to the medium but the *observer is moving* through it is

$$\nu' = \nu\left(\frac{v \pm v_o}{v}\right), \tag{20–9}$$

where the *plus* sign holds for motion *toward* the source and the *minus* sign holds for motion *away* from the source. Notice that the cause of the change here is the fact that the observer intercepts more or fewer waves each second because of his motion through the medium.

When the *source* is in motion *toward* a stationary observer, the effect is a shortening of the wavelength (see Fig. 20–11), for the source is following after the approaching waves and the crests therefore come closer together. If the frequency of the source is $\nu$ and its speed is $v_s$, then during each vibration it travels a distance $v_s/\nu$ and each wavelength is shortened by this amount. Hence, the wavelength of the sound arriving at the observer is not $\lambda = v/\nu$ but $\lambda' = v/\nu - v_s/\nu$. Therefore, the frequency of the sound heard by the observer is *increased*, being

$$\nu' = \frac{v}{\lambda'} = \frac{v}{(v - v_s)/\nu} = \nu\left(\frac{v}{v - v_s}\right). \tag{20–10a}$$

If the source moves *away* from the observer, the wavelength emitted is $v_s/\nu$ greater than $\lambda$, so that the observer hears a *decreased* frequency, namely

$$\nu' = \frac{v}{(v + v_s)/\nu} = \nu\left(\frac{v}{v + v_s}\right). \tag{20–10b}$$

Hence, the general relation holding when the *observer is at rest* with respect to the medium but the *source is moving* through it is

$$\nu' = \nu\left(\frac{v}{v \mp v_s}\right), \tag{20–10}$$

where the *minus* sign holds for motion *toward* the observer and the *plus* sign holds for motion *away* from the observer. Notice that the cause of the change here is the fact that the motion of the source through the medium shortens or increases the wavelength transmitted through the medium.

If both source *and* observer move through the transmitting medium, the student should be able to show that the observer hears a frequency

$$\nu' = \nu\left(\frac{v \pm v_o}{v \mp v_s}\right). \tag{20–11}$$

where the upper signs (+ numerator, − denominator) correspond to the source and observer moving along the line joining the two in the direction *toward* the other, and the lower signs in the direction *away* from the other. Notice that Eq. 20–11 reduces to Eq. 20–9 when $v_s = 0$ and to Eq. 20–10 when $v_o = 0$, as it must.

If a vibrating tuning fork on its resonating box is moved rapidly toward a wall, the observer will hear two notes of different frequency. One is the note heard directly from the receding fork and is lowered in pitch by the motion. The other note is due to the waves reflected from the wall, and this is raised in pitch. The superposition of these two wave trains produces beats.

The Doppler effect is important in light. The speed of light is so great that only astronomical or atomic sources, which have high velocities compared to terrestrial macroscopic sources, show pronounced Doppler effects. The astronomical effect consists of a shift in the wavelength observed from light emitted by elements on moving astronomical bodies compared to the wavelength observed from these same elements on earth. (See Chapter 40.) An easily observed consequence of the Doppler effect is the broadening (or spread in frequency) of the radiation emitted from hot gases. This broadening results from the fact that the emitting atoms or molecules move in all directions and with varying speeds relative to the observing instruments, so that a spread of frequencies is detected.

There are differences, however, in the Doppler effect formula for light and for sound. In sound it is not just the relative motion of source and observer that determines the frequency change. In fact, as we have seen, even when the relative motion is the same ($v_o$ in Eq. 20–9$a$ equals $v_s$ in Eq. 20–10$a$), we obtain different quantitative results, depending on whether the source or the observer is moving. This difference occurs because $v_o$ and $v_s$ are measured relative to the medium in which the sound wave is propagated and because this medium determines the wave speed. Light, however, does not require a material medium for its transmission, and the speed of light relative to the source or the observer is always the same value $c$, regardless of the motion of these bodies relative to each other. This is a basic postulate of the special theory of relativity. Hence, for light only the relative motion of source and observer can lead to physical changes, there being no material medium to use as a reference frame. Although the Doppler formula for light (Chapter 40) differs from that for sound, the effects are qualitatively the same. We can apply Eq. 20–10 to light as a good approximation if $v_s$ is taken to mean the *relative* velocity of source and observer and if $v_s$ is very small compared to the velocity of light.

▶ **Example 3.** Show that Eqs. 20–9 and 20–10 become practically identical when the speed of the sources and the observer are small compared to the speed of sound in the medium.

Let $v_o = v_s = u$. That is, let $u$ represent the speed of observer *or* source. Then Eq. 20–9 becomes

$$\nu' = \nu\left(1 \pm \frac{u}{v}\right).$$

We must show then that Eq. 20–10,

$$\nu' = \nu\left(\frac{v}{v \mp u}\right),$$

reduces to the previous form when $u/v \ll 1$.

We can rewrite Eq. 20–10 as

$$\nu' = \nu\left(\frac{1}{1 \mp u/v}\right).$$

Now by the binomial expansion

$$\left(\frac{1}{1 \mp u/v}\right) = \left(1 \mp \frac{u}{v}\right)^{-1} = 1 \pm \frac{u}{v} + \left(\frac{u}{v}\right)^2 \pm \cdots.$$

But if $u/v$ is sufficiently small compared to unity that we may neglect $(u/v)^2$ and higher powers, then

$$\left(\frac{1}{1 \mp u/v}\right) \cong 1 \pm \frac{u}{v},$$

and Eq. 20–10 becomes
$$\nu' \cong \nu\left(1 \pm \frac{u}{v}\right),$$

the same as Eq. 20–9.

As a numerical example take $u = 73.0$ miles/hr. The speed of sound in air is about 730 miles/hr. Then if the source has a speed $v_s = u = 73.0$ miles/hr toward the stationary observer, the frequency heard by the observer is Eq. 20–10,

$$\nu' = \nu\left(\frac{v}{v - v_s}\right) = \nu\left(\frac{730}{730 - 73.0}\right)$$

or

$$\frac{\nu'}{\nu} = 1.11.$$

If the observer has a speed $v_o = u = 73.0$ miles/hr toward the stationary source, the frequency heard by the observer is Eq. 20–9,

$$\nu' = \nu\left(\frac{v + v_o}{v}\right) = \nu\left(\frac{730 + 73.0}{730}\right).$$

or

$$\frac{\nu'}{\nu} = 1.10.$$

Hence, when $u/v = 73.0/730 = 1/10$, the percentage difference in the frequency heard between that for the moving observer and that for the moving source, the relative motion being the same, is only 1%.

When $v_o$ or $v_s$ becomes comparable in magnitude to $v$, the formulas just given for the Doppler effect must be modified. The modification is required because the linear relation between restoring force and displacement assumed up until now no longer holds in the medium. The speed of wave propagation is no longer the normal phase velocity, and the wave shapes change in time. Components of the motion at right angles to the line joining source and observer also contribute to the Doppler effect at these high speeds. When $v_o$ or $v_s$ exceeds $v$, the Doppler formula clearly has no meaning.

There are many instances in which the source moves through a medium at a speed greater than the phase velocity of the wave in that medium. In such cases the wavefront takes the shape of a cone with the moving body at its apex. Some examples are the bow wave from a speedboat on the water and the "shock wave" from an airplane or projectile moving through the air at a speed greater than the

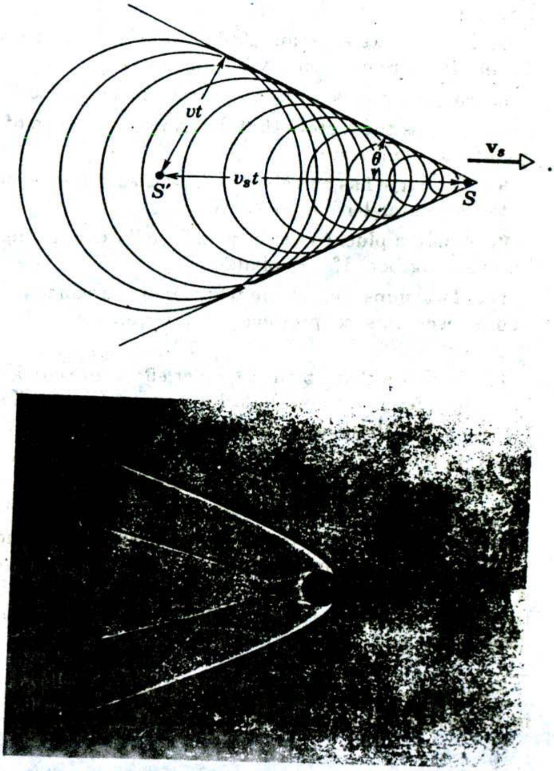


**Fig. 20–12** Above, right, a group of wavefronts associated with a projectile moving with supersonic speed. The wavefronts are spherical and their envelope is a cone. The student should see the relation between this figure and the previous one. Below, right, a spark photograph of a projectile undergoing this motion. (U.S. Navy Photograph.)

velocity of sound in that medium (supersonic speeds). The Cerenkov radiation consists of light waves emitted by charged particles which move through a medium with a speed greater than the phase velocity of light in that medium.*

In Fig. 20–12 we show the present positions of the spherical waves which originated at various positions of the source during its motion. The radius of each sphere at this time is the product of the wave speed $v$ and the time $t$ which has elapsed since the source was at its center. The envelope of these waves is a cone whose surface makes an angle $\theta$ with the direction of motion of the source. From the figure we obtain the result

$$\sin \theta = \frac{v}{v_s}.$$

For water waves the cone reduces to a pair of intersecting lines. In aerodynamics the ratio $v_s/v$ is called the Mach number.

## QUESTIONS

**1.** List some sources of infrasonic waves. Of ultrasonic waves.

**2.** What experimental evidence is there for assuming that the speed of sound is the same for all wavelengths?

**3.** What quantity, if any, for transverse waves in a string corresponds to the pressure amplitude for longitudinal waves in a tube?

**4.** A bell is rung for a short time in a school. After a while its sound is inaudible. Trace the sound waves and the energy they transfer from the time of emission until they become inaudible.

**5.** How can we experimentally locate the positions of nodes and antinodes in a string? In an air column? On a vibrating surface?

**6.** Explain how a stringed instrument is "tuned."

**7.** Discuss the factors that determine the range of frequencies in your voice and the quality of your voice.

**8.** The bugle has no valves. How then can we sound different notes on it? To what notes is the bugler limited? Why?

**9.** Would a plucked violin string oscillate for a longer or shorter time if the violin had no sounding board? Explain.

**10.** Two ships with steam whistles of the same pitch sound off in the harbor. Would you expect this to produce an interference pattern with regions of high and low intensity?

**11.** Suppose that, in the Doppler effect for sound, the source and receiver are at rest in some reference frame but the transmitting medium is moving with respect to this frame. Will there be a change in wavelength, or in frequency, received?

**12.** Is there a Doppler effect for sound when the observer or the source moves at right angles to the line joining them? How then can we determine the Doppler effect when the motion has a component at right angles to this line?

**13.** A satellite emits radio waves of constant frequency. These waves are picked up on the ground and made to beat against some standard frequency. The beat frequency is then sent through a loudspeaker and one "hears" the satellite signals. Describe how the sound changes as the satellite approaches, passes overhead, and recedes from the detector on the ground.

---

* See "Cerenkov Radiation: its Origin, Properties and Applications," by J. V. Jelley in *Contemporary Physics*, October 1961.

14. Two identical tuning forks emit notes of the same frequency. Explain how you might hear beats between them.

15. Transverse waves in a string can be polarized (see, for example, Question 16 of Chapter 19). Can sound waves be polarized?

## PROBLEMS

1. The lowest pitch detectable as sound by the average human ear consists of about 20 vib/sec and the highest of about 20,000 vib/sec. What is the wavelength of each in air?

2. A sound wave has a frequency of 440 vib/sec. What is the wavelength of this sound in air? In water?

3. Bats emit ultrasonic waves. The shortest wavelength emitted in air by a bat is about 0.13 in. What is the highest frequency a bat can emit?

4. (*a*) A loudspeaker has a diameter of 6.0 in. At what frequency will the wavelength of the sound it emits in air be equal to its diameter? Be ten times its diameter? Be one-tenth its diameter? (*b*) Make the same calculations for a speaker of diameter 12.0 in. If the wavelength is large compared to the diameter of the speaker, the sound waves spread out almost uniformly in all directions from the speaker, but when the wavelength is small compared to the diameter of the speaker, the wave energy is propagated mostly in the forward direction.

5. A rule for finding your distance from a lightning flash is to count seconds from the time you see the flash until you hear the thunder and then divide the count by five. The result is supposed to give the distance in miles. Explain this rule and determine the per cent error in it at standard conditions.

6. A stone is dropped into a well. The sound of the splash is heard at a time $t$ later. What is the depth $d$ of the well? Find $d$ when $t = 3.0$ sec.

7. (*a*) The speed of sound in a certain metal is $V$. One end of a pipe of that metal of length $l$ is struck a blow. A listener at the other end hears two sounds, one from the wave that has traveled along the pipe and the other from the wave that has traveled through the air. If $v$ is the speed of sound in air, what time interval $t$ elapses between the two sounds? (*b*) Suppose $t = 1.4$ sec and the metal is iron. Find the length $l$.

8. The pressure in a traveling sound wave is given by the equation

$$p = 1.5 \sin \pi(x - 330t),$$

where $x$ is in meters, $t$ in seconds, and $p$ in nt/meter$^2$. Find the pressure amplitude, frequency, wavelength, and speed of the wave.

9. Show that the intensity of a sound wave (*a*) when expressed in terms of the pressure amplitude $P$, is given by

$$I = \frac{P^2}{2\rho_0 v},$$

where $v$ is the speed of the wave and $\rho_0$ is the standard density of air, and, (*b*) when expressed in terms of the displacement amplitude $y_m$, is given by

$$I = 2\pi^2 \rho_0 v y_m{}^2 \nu^2,$$

where $\nu$ is the frequency of the wave.

10. (*a*) If two sound waves, one in air and one in water, are equal in intensity, what is the ratio of the pressure amplitude of the wave in water to that of the wave in air? (*b*) If the pressure amplitudes are equal instead, what is the ratio of the intensities of the waves?

**11.** A note of frequency 300 vib/sec has an intensity of 1.0 microwatt/meter$^2$. What is the amplitude of the air vibrations caused by this sound?

**12.** Two waves give rise to pressure variations at a certain point in space given by

$$p_1 = P \sin 2\pi\nu t,$$

$$p_2 = P \sin 2\pi(\nu t - \phi).$$

What is the amplitude of the resultant wave at this point when $\phi = 0$, $\phi = \frac{1}{4}$, $\phi = \frac{1}{6}$, $\phi = \frac{1}{8}$?

**13.** In Fig. 20–13 we show an acoustic interferometer, used to demonstrate the interference of sound waves. $S$ is a diaphragm that vibrates under the influence of an electromagnet. $D$ is a sound detector, such as the ear or a microphone. Path $SBD$ can be varied in length, but path $SAD$ is fixed.

The interferometer contains air, and it is found that the sound intensity has a minimum value of 100 units at one position of $B$ and continuously climbs to a maximum value of 900 units at a second position 1.65 cm from the first. Find (*a*) the frequency of the sound emitted from the source, and (*b*) the relative amplitudes of the two waves arriving at the detector. (*c*) How can it happen that these waves have different amplitudes, considering that they originate at the same source?

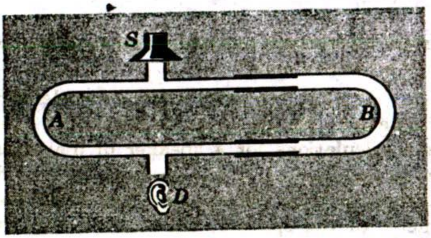


**Fig. 20–13**

**14.** Two loudspeakers, $S_1$ and $S_2$, each emit sound of frequency 200 vib/sec uniformly in all directions. $S_1$ has an acoustic output of $1.2 \times 10^{-3}$ watt and $S_2$ one of $1.8 \times 10^{-3}$ watt. $S_1$ and $S_2$ vibrate in phase. Consider a point $P$ which is 4.0 meters from $S_1$ and 3.0 meters from $S_2$. (*a*) How are the phases of the two waves arriving at $P$ related? (*b*) What is the intensity of sound at $P$ with both $S_1$ and $S_2$ on? (*c*) What is the intensity of sound at $P$ if $S_1$ is turned off ($S_2$ on)? (*d*) What is the intensity of sound at $P$ if $S_2$ is turned off ($S_1$ on)?

**15.** A spherical sound source is placed at $P_1$ near a reflecting wall $AB$ and a microphone is located at point $P_2$, as shown in Fig. 20–14. The frequency of the sound source $P_1$ is variable. Find two different frequencies for which the sound intensity, as observed at $P_2$, will be a maximum. The speed of sound in air is 1100 ft/sec.

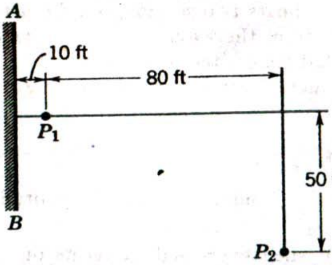


**Fig. 20–14**

**16.** The water level in a vertical glass tube 1.0 meter long can be adjusted to any position in the tube. A tuning fork vibrating at 660 vib/sec is held just over the open top end of the tube. At what positions of the water level will there be resonance?

**17.** In Fig. 20–15 a rod $R$ is clamped at its center and a disk $D$ at its end projects into a glass tube, which has cork filings spread over its interior. A plunger $P$ is provided at the other end of the tube. The rod is set into longitudinal vibration and the plunger is moved until the filings form a pattern of nodes and antinodes (the filings form well-defined ridges at the antinodes). If we know the frequency $\nu$ of the longitudinal vibrations in the rod, a measurement of the average distance $d$ between successive antinodes

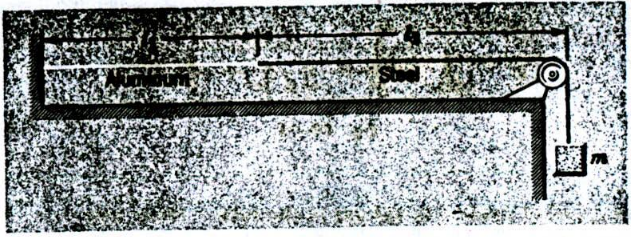


Fig. 20–17

**26.** Two identical piano wires have a fundamental frequency of 600 vib/sec when kept under the same tension. What fractional increase in the tension of one wire will lead to the occurrence of six beats per second when both wires vibrate simultaneously?

**27.** A tuning fork of unknown frequency makes three beats per second with a standard fork of frequency 384 vib/sec. The beat frequency decreases when a small piece of wax is put on a prong of the first fork. What is the frequency of this fork?

**28.** Microwaves, which travel with the speed of light, are reflected from a distant airplane approaching the wave source. It is found that when the reflected waves are beat against the waves radiating from the source the beat frequency is 990 cycles/sec. If the microwaves are 0.10 meter in wavelength, what is the approach speed of the airplane?

**29.** Could you go through a red light fast enough to have it appear green? Would you get a ticket for speeding? Take $\lambda = 6200 \times 10^{-8}$ cm for red light, $\lambda = 5400 \times 10^{-8}$ cm for green light, and $c = 3 \times 10^{10}$ cm/sec as the speed of light.

**30.** A whistle of frequency 540 vib/sec rotates in a circle of radius 2.00 ft at an angular speed of 15.0 radians/sec. What is the lowest and the highest frequency heard by a listener a long distance away at rest with respect to the center of the circle?

**31.** A siren emitting a sound of frequency 1000 vib/sec moves away from you toward a cliff at a speed of 10 meters/sec. (*a*) What is the frequency of the sound you hear coming directly from the siren? (*b*) What is the frequency of the sound you hear reflected off the cliff? (*c*) What beat frequency would you hear? Take the speed of sound in air as 330 meters/sec.

**32.** Trooper *B* is chasing speeder *A* along a straight stretch of road. Both are moving at a top speed of about 100 miles/hr, which is about 150 ft/sec. Trooper *B*, failing to catch up, sounds his siren again. Take the speed of sound in air to be 1100 ft/sec and the frequency of the source to be 500 cycles/sec. Demonstrate clearly whether there will be a Doppler shift in the frequency heard by speeder *A* and, if there is, what the frequency change is.

**33.** A source of sound waves of frequency 1080 vib/sec moves to the right with a speed of 108 ft/sec relative to the ground. To its right is a reflecting surface moving to the left with a speed of 216 ft/sec relative to the ground. Take the speed of sound in air to be 1080 ft/sec and find (*a*) the wavelength of the sound emitted in air by the source, (*b*) the number of waves per second arriving at the reflecting surface, (*c*) the speed of the reflected waves, (*d*) the wavelength of the reflected waves.

**34.** A bullet is fired with a speed of 2200 ft/sec. Find the angle made by the shock wave with the line of motion of the bullet.

**35.** A jet plane passes overhead at a height of 5000 meters and a speed of Mach 1.5 (that is, 1.5 times the speed of sound). (*a*) Find the angle made by the shock wave with

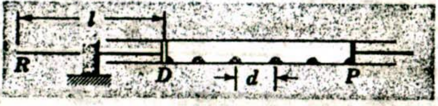


Fig. 20–15

determines the speed of sound $v$ in the gas in the tube. Show that

$$v = 2\nu d.$$

This is Kundt's method for determining the speed of sound in various gases.

**18.** A tube 1.0 meter long is closed at one end. A stretched wire is placed near the open end. The wire is 0.30 meter long and has a mass of 0.010 kg. It is held fixed at both ends and vibrates in its fundamental mode. It sets the air column in the tube into vibration at its fundamental frequency by resonance. Find (*a*) the frequency of oscillation of the air column and (*b*) the tension in the wire.

**19.** A tube can act like an acoustic filter, discriminating against the passage through it of sound of frequencies different from the natural frequencies of the tube. The muffler of an automobile is an example. (*a*) Explain how such a filter works. (*b*) How can we determine the cut-off frequency, below which frequency sound is not transmitted?

**20.** An open organ pipe has a fundamental frequency of 300 vib/sec. The first overtone of a closed organ pipe has the same frequency as the first overtone of the open pipe. How long is each pipe?

**21.** $S$ in Fig. 20–16 is a small loudspeaker driven by an audio oscillator and amplifier, adjustable in frequency from 1000 to 2000 cycles/sec only. $D$ is a piece of cylindrical sheet-metal pipe 18.0 in. long. If the velocity of sound in air is 1130 ft/sec at the existing temperature, at what frequencies will resonance occur when the frequency emitted by the speaker is varied from 1000 to 2000 cycles/sec? Sketch the displacement modes for each. Neglect end effects.

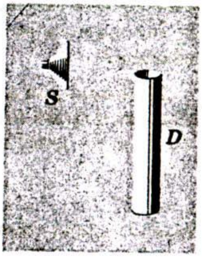


Fig. 20–16

**22.** A certain violin string is 50 cm long between its fixed points and has a mass of 2.0 gm. The string sounds an $A$ note (440 vib/sec) when played without fingering. Where must one put one's finger to play a $C$ (528 vib/sec)?

**23.** The strings of a cello have a length $L$. By what length $l$ must they be shortened by fingering to change the pitch by a frequency ratio $r$? Find $l$, if $L = 0.80$ meter and $r = 6/5, \ldots, r = 3/2$.

**24.** If a violin string is tuned to a certain note, by how much must the tension in the string be increased if it is to emit a note of double the original frequency (that is, a note one octave higher in pitch).

**25.** An aluminum wire of length $l_1 = 60.0$ cm and of cross-sectional area $1.00 \times 10^{-2}$ cm$^2$ is connected to a steel wire of the same cross-sectional area. The compound wire, loaded with a block $m$ of mass 10.0 kg, is arranged as shown in Fig. 20–17, so that the distance $l_2$ from the joint to the supporting pulley is 86.6 cm. Transverse waves are set up in the wire by using an external source of variable frequency. (*a*) Find the lowest frequency of excitation for which standing waves are observed such that the joint in the wire is a node. (*b*) What is the total number of nodes observed at this frequency, excluding the two at the ends of the wire? The density of aluminum is 2.60 gm/cm$^3$, and that of steel is 7.80 gm/cm$^3$.

the line of motion of the jet. (*b*) How long after the jet has passed directly overhead will the shock wave reach the ground?

**36.** The speed of light in water is about three-fourths the speed of light in vacuum. A beam of high-speed electrons from a betatron emits Cerenkov radiation in water, the wavefront being a cone of angle 60°. Find the speed of the electrons in the water.

**37.** Calculate the speed of the projectile illustrated in the photograph in Fig. 20–12. Assume the speed of sound in the medium through which the projectile is traveling to be 380 meters/sec.

# Temperature

## CHAPTER 21

### 21–1 Macroscopic and Microscopic Descriptions

In analyzing physical situations we usually focus our attention on some portion of matter which we separate, in our minds, from the environment external to it. We call such a portion the *system*. Everything outside the system which has a direct bearing on its behavior we call the *environment*. We then seek to determine the behavior of the system by finding how it interacts with its environment. For example, a ball can be the system and the environment can be the air and the earth. In free fall we seek to find how the air and the earth affect the motion of the ball. Or the gas in a container can be the system, and a movable piston and a Bunsen burner can be the environment. We seek to find how the behavior of the gas is affected by the action of the piston and burner. In all such cases we must choose suitable observable quantities to describe the behavior of the system. We classify these quantities, which are gross properties of the system measured by laboratory operations, as *macroscopic*. For processes in which heat is involved the laws relating the appropriate macroscopic quantities (which include pressure, volume, temperature, internal energy, and entropy, among others) form the basis for the science of *thermodynamics*. Many of the macroscopic quantities (pressure, volume, and temperature, for example) are directly associated with our sense perceptions. We can also adopt a *microscopic* point of view. Here we consider quantities that describe the atoms and molecules that make up the system, their speeds, energies, masses, angular momenta, behavior during collisions, etc. These quantities, or mathematical formulations based on them, form the basis for the science of *statistical mechanics.* The microscopic properties are not directly associated with our sense perceptions.

For any system the macroscopic and the microscopic quantities must be related because they are simply different ways of describing the same situation. In particular, we should be able to express the former in terms of the latter. The pressure of a gas, viewed macroscopically, is measured operationally using a manometer (Fig. 17-10). Viewed microscopically it is related to the average rate per unit area at which the molecules of the gas deliver momentum to the manometer fluid as they strike its surface. In Section 23-4 we will make this microscopic definition of pressure quantitative. Similarly (see Section 23-5), the temperature of a gas may be related to the average kinetic energy of translation of the molecules.

If the macroscopic quantities can be expressed in terms of the microscopic quantities, we should be able to express the laws of thermodynamics quantitatively in the language of statistical mechanics. We can indeed do this. In the words of R. C. Tolman:

> The explanation of the complete science of thermodynamics in terms of the more abstract science of statistical mechanics is one of the greatest achievements of physics. In addition, the more fundamental character of statistical mechanical considerations makes it possible to supplement the ordinary principles of thermodynamics to an important extent.

We begin our examination of heat phenomena in this chapter with a study of temperature. As we progress we shall try to gain a deeper understanding of these phenomena by interweaving the microscopic and the macroscopic description—statistical mechanics and thermodynamics. The interweaving of the microscopic and the macroscopic points of view is characteristic of modern physics.

## 21-2 Thermal Equilibrium—The Zeroth Law of Thermodynamics

The sense of touch is the simplest way to distinguish hot bodies from cold bodies. By touch we can arrange bodies in the order of their hotness, deciding that $A$ is hotter than $B$, $B$ than $C$, etc. We speak of this as our *temperature* sense. This is a very subjective procedure for determining the temperature of a body and certainly not very useful for purposes of science. A simple experiment, suggested in 1690 by John Locke, shows the unreliability of this method. Let a person immerse his hands, one in hot water, the other in cold. Then let him put both hands in water of intermediate hotness. This will seem cooler to the first hand and warmer to the second hand. Our judgment of temperature can be rather misleading. Further, the range of our temperature sense is limited. What we need is an objective, numerical, measure of temperature.

To begin with, we should try to understand the meaning of temperature. Let an object $A$ which feels cold to the hand and an identical object $B$ which feels hot be placed in contact with each other. After a sufficient length of time, $A$ and $B$ give rise to the same temperature sensation.

Then $A$ and $B$ are said to be in *thermal equilibrium* with each other. We can generalize the expression "two bodies are in thermal equilibrium" to mean that the two bodies are in states such that, if the two *were* connected, the combined systems would be in thermal equilibrium. The logical and operational test for thermal equilibrium is to use a third or test body, such as a thermometer. This is summarized in a postulate often called *the zeroth law of thermodynamics: If A and B are in thermal equilibrium with a third body C (the "thermometer"), then A and B are in thermal equilibrium with each other*

This discussion expresses the idea that the temperature of a system is a property which eventually attains the same value as that of other systems when all these systems are put in contact. This concept agrees with the everyday idea of temperature as the measure of the hotness or coldness of a system, because as far as our temperature sense can be trusted, the hotness of all objects becomes the same after they have been in contact long enough. The idea contained in the zeroth law, although simple, is not obvious. For example, Jones and Smith each know Green, but they may or may not know each other. Two pieces of iron attract a magnet but they may or may not attract each other.

A more formal, but perhaps more fundamental phrasing of the zeroth law is: *There exists a scalar quantity called* temperature, *which is a property of all thermodynamic systems (in equilibrium states), such that temperature equality is a necessary and sufficient condition for thermal equilibrium.* This statement* justifies our use of temperature as a thermodynamic variable; the formulation given above is the corollary of this new statement. Speaking loosely, the essence of the zeroth law is: *there exists a useful quantity called "temperature."*

### 21–3 Measuring Temperature

There are many measurable physical properties that vary as our physiological perception of temperature varies. Among these are the volume of a liquid, the length of a rod, the electrical resistance of a wire, the pressure of a gas kept at constant volume, the volume of a gas kept at constant pressure, and the color of a lamp filament. Any of these properties can be used in the construction of a thermometer—that is, in the setting up of a particular "private" temperature scale. Such a temperature scale is established by choosing a particular thermometric substance and a particular thermometric property of this substance. We then define the temperature scale by an *assumed* continuous monotonic relation between the chosen thermometric property of our substance and the temperature as measured on our (private) scale. For example, the thermometric substance may be a liquid in a glass capillary tube and the thermometric property can be the length of the liquid column; or the thermometric substance may be a gas kept in a container at constant volume and the thermometric property can be the pressure of the gas; and so forth. *We must realize that each choice of thermometric substance and property—along with the assumed rela-*

---

* See J. S. Thomsen, *American Journal of Physics*, 30, 294, 1962.

*tion between property and temperature—leads to an individual temperature scale whose measurements need not necessarily agree with measurements made on any other independently defined temperature scale.*

This apparent chaos in the definition of temperature is removed by universal agreement, within the scientific community, on the use of a particular thermometric substance, a particular thermometric property, and a particular functional relation between measurements of that property and a universally accepted temperature scale. A private temperature scale defined in any other way can then always be calibrated against the universal scale. We describe such a universal scale in Section 21–5 and an equivalent one in Section 25–6.

Suppose that we have chosen a thermometric substance. Let us represent by $X$ the thermometric property that we wish to use in setting up a temperature scale. We arbitrarily choose the following linear function of the property $X$ as the temperature $T$ which the appropriate thermometer, and any system in thermal equilibrium with it, has:

$$T(X) = aX. \tag{21–1}$$

In this expression $a$ is a constant which we must still evaluate. By choosing this linear form for $T(X)$ we have fixed it so that *equal temperature differences*, or temperature intervals, *correspond to equal changes in* $X$. This means, for example, that every time the mercury column in the mercury-in-glass thermometer changes in length by one unit, the temperature changes by a definite fixed amount, no matter what the starting temperature. It also follows that two temperatures, measured with the same thermometer, are in the same ratio as their corresponding $X$'s, that is

$$\frac{T(X_1)}{T(X_2)} = \frac{X_1}{X_2}.$$

To determine the constant $a$, and hence to calibrate the thermometer, we specify a *standard fixed point* at which all thermometers must give the same reading for temperature $T$. This fixed point is chosen to be that at which ice, liquid water, and water vapor coexist in equilibrium and is called the *triple point of water*. This state can be achieved only at a definite pressure and is unique (Fig. 21–1). The water vapor pressure at the triple point is 4.58 mm-Hg. The temperature at this standard fixed point is arbitrarily* set at 273.16 degrees Kelvin and is abbreviated as 273.16° K. The Kelvin degree is a unit temperature interval.

If we indicate values at the triple point by the subscript *tr*, then, for any thermometer,

$$\frac{T(X)}{T(X_{tr})} = \frac{X}{X_{tr}},$$

---

* **Adopted in 1954 at the Tenth General Conference on Weights and Measures in Paris.**

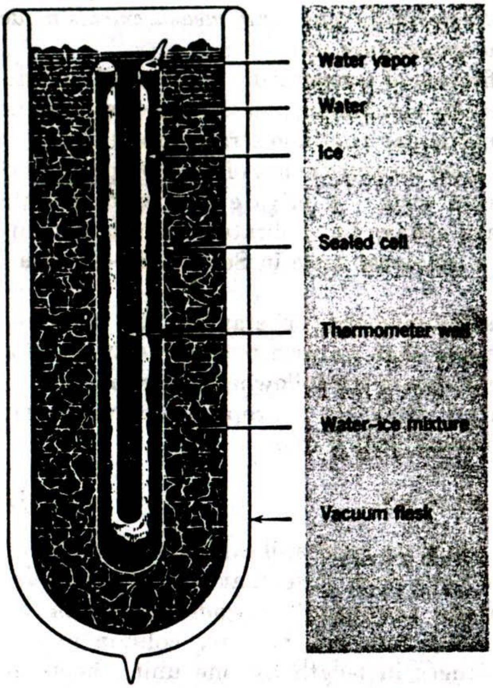


**Fig. 21–1** The National Bureau of Standards triple-point cell. It contains pure water and is sealed after all air has been removed. It is then immersed in a water-ice bath. The system is at the triple point when ice, water, and vapor are all present, and in equilibrium, inside the cell. The thermometer to be calibrated is immersed in the central well.

where, for *all* thermometers,

$$T(X_{tr}) = 273.16° \text{ K},$$

so that

$$T(X) = 273.16° \text{ K} \frac{X}{X_{tr}}. \tag{21–2}$$

Hence, when the thermometric property has the value $X$, the temperature $T$, on the particular private scale selected, is given in °K by $T(X)$, when the value of $X$ and $X_{tr}$ are inserted on the right-hand side of this equation.

We can now apply Eq. 21–2 to several thermometers. For a liquid-in glass thermometer $X$ is $L$, the length of the liquid column, and Eq. 21–2 yields

$$T(L) = 273.16° \text{ K} \frac{L}{L_{tr}}.$$

For a gas at constant pressure, $X$ is $V$, the volume of the gas, and

$$T(V) = 273.16° \text{ K} \frac{V}{V_{tr}} \qquad \text{(constant } P\text{)}.$$

For a gas at constant volume, $X$ is $P$, the gas pressure, and

$$T(P) = 273.16° \text{ K} \frac{P}{P_{tr}} \qquad \text{(constant } V\text{)}.$$

For a platinum resistance thermometer, $X$ is $R$, the electrical resistance, and

$$T(R) = 273.16^\circ \text{ K} \frac{R}{R_{tr}},$$

and likewise for other thermometric substances and thermometric properties.

▶ **Example 1.** A certain platinum resistance thermometer has a resistance $R$ of 90.35 ohms when its bulb is placed in a triple-point cell like that of Fig. 21-1. What temperature is defined by Eq. 21-2 if the bulb is placed in an environment such that its resistance is 96.28 ohms?

From Eq. 21-2,

$$T(X) = 273.16^\circ \text{ K} \frac{X}{X_{tr}}$$

$$= (273.16^\circ \text{ K}) \left(\frac{96.28}{90.35}\right) = 280.6^\circ \text{ K}.$$

Note that this temperature is on a private scale, defined by applying Eq. 21-2 to a particular device, the platinum resistance thermometer. ◀

The question now arises whether the value we obtain for the temperature of a system depends on the choice of the thermometer we use to measure it. We have insured by definition that all the different kinds of thermometers will agree at the standard fixed point, but what happens at other points? We can imagine a series of tests in which the temperature of a given system is measured simultaneously with many different thermometers. Results of such tests show that the thermometers all read differently. Even when different thermometers of the same kind are used, such as constant-volume gas thermometers using different gases, we obtain different temperature readings for a given system in a given state.

Hence, to obtain a definite temperature scale, we must select one particular kind of thermometer as the standard. The choice will be made, not on the basis of experimental convenience, but by inquiring whether the temperature scale defined by a particular thermometer proves to be a useful quantity in the formulation of the laws of physics. The smallest variation in readings is found among different constant-volume gas thermometers, which suggests that we choose a gas as the standard thermometric substance. It turns out that as the amount of gas used in such a thermometer, and therefore its pressure, is reduced, the variation in readings between gas thermometers using different kinds of gas is reduced also. Hence, there seems to be something fundamental about the behavior of a constant-volume thermometer containing a gas at low pressure.

## 21-4 The Constant Volume Gas Thermometer

If the volume of a gas is kept constant, its pressure depends on the temperature and increases steadily with rising temperature. The constant-volume gas thermometer uses the pressure at constant volume as the thermometric property.

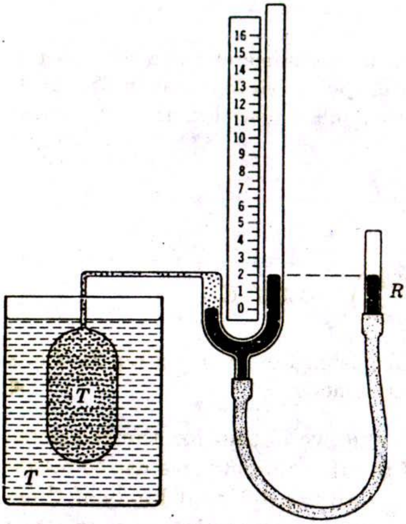


**Fig. 21–2** A representation of a constant-volume gas thermometer. As long as the mercury in the left manometer tube remains at the same position on the scale (zero) the volume of the confined gas will be constant. The meniscus can always be brought to the zero position by raising or lowering reservoir *R*.

The thermometer is shown diagrammatically in Fig. 21–2. It consists of a bulb of glass, porcelain, quartz, platinum or platinum-iridium (depending on the temperature range over which it is to be used), connected by a capillary tube to a mercury manometer. The bulb containing some gas is put into the bath or environment whose temperature is to be measured; by raising or lowering the mercury reservoir the mercury in the left branch of the U-tube can be made to coincide with a fixed reference mark, thus keeping the confined gas at a constant volume. Then we read the height of the mercury in the right branch. The pressure of the confined gas is the difference of the heights of the mercury columns (times $\rho g$) plus the atmospheric pressure, as indicated by the barometer. In practice the apparatus is very elaborate and we must make many corrections, for example, (1) to allow for the small volume change owing to slight contraction or expansion of the bulb and (2) to allow for the fact that not all the confined gas (such as that in the capillary) has been immersed in the bath. Assume that all corrections have been made, and let $P$ be the corrected value of the pressure at the temperature of the bath. Then the temperature is given provisionally (see below) by

$$T(P) = 273.16^\circ \text{ K} \frac{P}{P_{tr}} \qquad \text{(constant } V\text{)}. \qquad (21\text{–}3)$$

The constant-volume thermometer, used as described below, is the thermometer which serves to establish the temperature scale used universally in scientific work today.

## 21–5 Ideal Gas Temperature Scale

Let a certain amount of gas be put into the bulb of a constant-volume gas thermometer so that when the bulb is surrounded by water at the triple point the pressure $P_{tr}$ is equal to a definite value, say 80 cm-Hg. Now surround the bulb with steam condensing at 1-atm pressure and, with the volume kept constant at its previous value, measure the gas pressure $P_s$, the pressure at the steam point, in this case, $P_{s80}$. Then calculate the temperature provisionally from $T(P_{s80}) = 273.16° \text{ K } (P_{s80}/80 \text{ cm-Hg})$. Next remove some of the gas so that $P_{tr}$ has a smaller value, say 40 cm-Hg. Then measure the new value of $P_s$ and calculate another provisional temperature from $T(P_{s40}) = 273.16° \text{ K } (P_{s40}/40 \text{ cm-Hg})$. Continue this same procedure, reducing the amount of gas in the bulb again, and at this new lower value of $P_{tr}$ calculating the temperature at the steam point $T(P_s)$. If we plot the values $T(P_s)$ against $P_{tr}$ and have enough data, we can extrapolate the resulting curve to the intersection with the axis where $P_{tr} = 0$.

In Fig. 21–3, we plot curves obtained from such a procedure for constant-volume thermometers of some different gases. These curves show that the temperature readings of a constant-volume gas thermometer depend on the gas used at ordinary values of the reference pressure. However, as

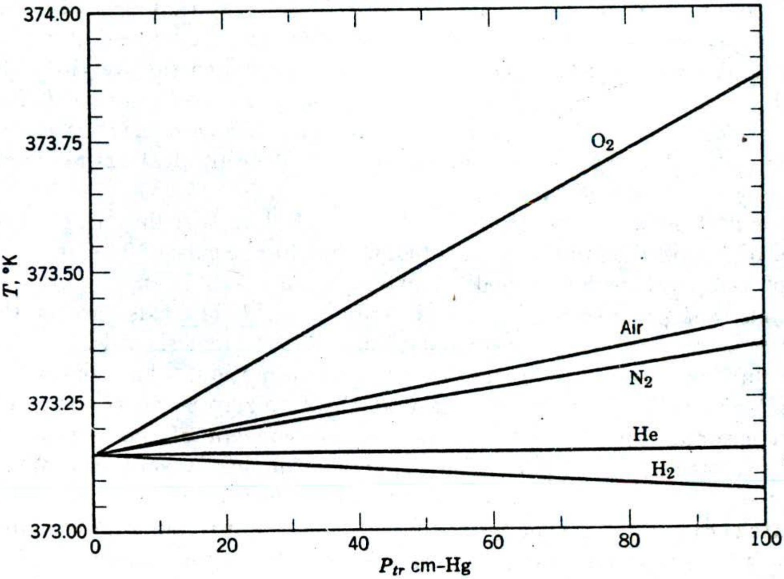


**Fig. 21–3** The readings of a constant-volume gas thermometer for the temperature $T$ of condensing steam as a function of $P_{tr}$, when different gases are used. As the amount of gas in the thermometer is reduced its pressure $P_{tr}$ at the triple point decreases. Note that at a particular $P_{tr}$ the values of $T$ given by different gas thermometers differ. The discrepancy is small but measurable, being about 0.2 per cent in the most extreme cases shown ($O_2$ and $H_2$ at 100 cm-Hg). Helium gives nearly the same $T$ at all pressures (the curve is almost horizontal) so that its behaviour is the most similar to that of an ideal gas over the entire range shown.

the reference pressure is decreased, the temperature readings of constant-volume gas thermometers using different gases approach the same value. Therefore, *the extrapolated value of the temperature depends only on the general properties of gases and not on any particular gas.* We therefore define an *ideal gas temperature scale* by the relation

$$T = 273.16^\circ\ \mathrm{K} \lim_{P_{tr}\to 0} \left(\frac{P}{P_{tr}}\right) \qquad \text{(constant } V\text{).} \tag{21–4}$$

Our standard thermometer is therefore chosen to be a constant-volume gas thermometer using a temperature scale defined by Eq. 21–4.

Although our temperature scale is independent of the properties of any one particular gas, it does depend on the properties of gases in general (that is, on the properties of an ideal gas). Therefore, to measure a temperature, a gas must be used at that temperature. The lowest temperature that can be measured with any gas thermometer is about 1° K. To obtain this temperature we must use low-pressure helium, for helium becomes a liquid at a temperature lower than any other gas. Therefore we cannot give experimental meaning to temperatures below about 1° K, by means of a gas thermometer.

We would like to define a temperature scale in a way that is *independent of the properties of any particular substance.* We will show in Section 25–6 that the *absolute thermodynamic temperature scale*, called the Kelvin scale, is such a scale. We will show also that *the ideal gas scale and the Kelvin scale are identical in the range of temperatures in which a gas thermometer may be used.* For this reason we can write °K after an ideal gas temperature, as we have already done.

We will also show in Section 25–6 that the Kelvin scale has an *absolute zero* of 0° K and that temperatures below this do not exist. The absolute zero of temperature has defied all attempts to reach it experimentally, although it is possible to come arbitrarily close.* The existence of the absolute zero is inferred by extrapolation. The student should not think of absolute zero as a state of zero energy and no motion. The conception that all molecular action would cease at absolute zero is incorrect. This notion assumes that the purely macroscopic concept of temperature is strictly connected to the microscopic concept of molecular motion. When we try to make such a connection, we find in fact that as we approach absolute zero the kinetic energy of the molecules approaches a finite value, the so-called zero-point energy. The molecular energy is a minimum, but not zero, at absolute zero.

---

* It is possible to prepare systems that have *negative Kelvin temperatures.* Surprisingly enough, such temperatures are *not* reached by passing through 0° K but by proceeding through infinite temperatures. That is, negative temperatures are not 'colder' than absolute zero but instead are 'hotter' than infinite temperatures. See *Science by Degrees*, by Castle, Emmerich, Heikes, Miller, and Rayne, published by Walker and Company, New York, 1965. The absolute zero remains experimentally unattainable.

**Table 21–1**

SOME TEMPERATURES* (°K)

| | |
|---|---|
| Carbon thermonuclear reaction | $5 \times 10^8$ |
| Helium thermonuclear reaction | $10^8$ |
| Solar interior | $10^7$ |
| Solar corona | $10^6$ |
| Shock wave in air at Mach 20 | $2.5 \times 10^4$ |
| Luminous nebulae | $10^4$ |
| Solar surface | $6 \times 10^3$ |
| Tungsten melts | $3.6 \times 10^3$ |
| Lead melts | $6.0 \times 10^2$ |
| Water freezes | $2.7 \times 10^2$ |
| Oxygen boils (1 atm) | $9.0 \times 10^1$ |
| Hydrogen boils (1 atm) | $2.0 \times 10^1$ |
| Helium ($He^4$) boils at 1 atm | 4.2 |
| $He^3$ boils at attainable low pressure | $3.0 \times 10^{-1}$ |
| Adiabatic demagnetization of paramagnetic salts | $10^{-3}$ |
| Adiabatic demagnetization of nuclei | $10^{-6}$ |

* See *Scientific American*, September 1954; special issue on heat.

In Table 21–1 we list the temperatures, on the Kelvin scale, of various bodies and processes.

## 21–6 The Celsius and Fahrenheit Scales

Two temperature scales in common use are the Celsius* and the Fahrenheit scales. These are defined in terms of the Kelvin scale, which is the fundamental temperature scale in science.

The Celsius temperature scale uses a degree (the unit of temperature) which has the same magnitude as the degree on the Kelvin scale. If we let $t$ represent the Celsius temperature, then

$$t = T - 273.15° \tag{21–5}$$

relates the Celsius temperature $t$ (°C) and the Kelvin temperature $T$ (°K). We see that the triple point of water (= 273.16° K by definition) corresponds to 0.01° C. By experiment the temperature at which ice and air-saturated water are in equilibrium at atmospheric pressure—the so-called ice point—proves to be 0.00° C and the temperature at which steam and liquid water are in equilibrium at 1-atm pressure—the so-called steam point—proves to be 100.00° C.

The Fahrenheit scale, in common use in English-speaking countries (except in England itself, which adopted the Celsius scale for commercial

---

* This scale, based on a scale invented by a Swede named Celsius in 1742, was called the "centigrade" scale until 1948, when the Ninth General Conference on Weights and Measures decided that the name should be changed.

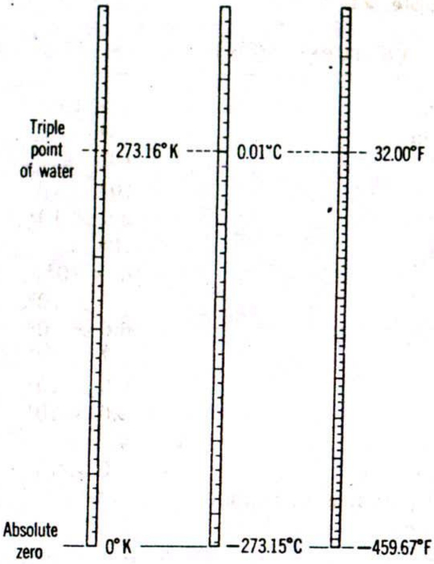


**Fig. 21–4** The Kelvin, Celsius, and Fahrenheit temperature scales.

and civil use in 1964) is not used in scientific work. The relationship between the Fahrenheit and Celsius scales is defined to be

$$T_F = 32^\circ \text{ F} + \tfrac{9}{5} T_C.$$

From this relation we can conclude that the ice point (0.00° C) equals 32.0° F, that the steam point (100.0° C) equals 212.0° F, and that one Fahrenheit degree is exactly $\frac{5}{9}$ as large as one Celsius degree. In Fig. 21–4 we compare the Kelvin, Celsius, and Fahrenheit scales.

## 21–7 The International Practical Temperature Scale

Let us now summarize the ideas of the last few sections. The standard fixed point in thermometry is the triple point of water which is arbitrarily assigned a value 273.16° K. The constant-volume gas thermometer is the standard thermometer. The extrapolated gas scale is used to define the ideal gas temperature from $T = 273.16^\circ \text{ K} \lim_{P_{tr}\to 0} (P/P_{tr})$. This scale is identical with the (absolute thermodynamic) Kelvin scale in the range in which a gas thermometer can be used.

By using the standard thermometer in this way, we can experimentally determine other reference points for temperature measurements, called fixed points. We list the basic fixed points adopted for experimental reference in Table 21–2. The temperatures can be expressed on the Celsius scale, with the use of Eq. 21–5, once the Kelvin temperature is determined.

Determining ideal gas temperatures is a painstaking job. It would not make sense to use this procedure to determine temperatures for all work.

Table 21-2

FIXED POINTS ON THE INTERNATIONAL PRACTICAL TEMPERATURE SCALE (1960)*

| Substance | Designation | Temperature °C | °K |
|---|---|---|---|
| Oxygen | Normal boiling point | −182.97 | 90.18 |
| Water | Triple point | 0.01 | 273.16 |
| Water | Normal boiling point | 100.00 | 373.15 |
| Sulfur† | Normal boiling point | 444.60 | 717.75 |
| Silver | Normal melting point | 960.80 | 1233.95 |
| Gold | Normal melting point | 1063.00 | 1336.15 |

* All temperatures assumed exact for the purposes of establishing the scale.
† The normal melting point of zinc (419.505° C) may be substituted.

Hence, an International Practical Temperature Scale (IPTS) was adopted in 1927 (revised in 1948 and again in 1954 and 1960) to provide a scale that can be used easily for practical purposes, such as for calibration of industrial or scientific instruments. This scale consists of a set of recipes for providing in practice the best possible approximations to the Kelvin scale. A set of fixed points, the basic points in Table 21-2, is adopted, and a set of instruments is specified to be used in interpolating between these fixed points and in extrapolating beyond the highest fixed point. Formulas are specified for correcting the basic temperatures according to the barometer reading. The IPTS departs from the Kelvin scale at temperatures between the fixed points, but the difference is usually negligible. The IPTS has become the legal standard in nearly all countries.

## 21-8 Temperature Expansion

Common effects of temperature changes are changes in size and changes of state of materials. Let us consider changes of sizes which occur without changes of state. Consider a simple model of a crystalline solid. The atoms are held together in a regular array by forces of electrical origin. The forces between atoms are like those that would be exerted by a set of springs connecting the atoms, so that we can visualize the solid body as a microscopic bedspring (Fig. 21-5). These "springs" are quite stiff (Problem 38, Chapter 15), and there are about $10^{22}$ of them per cubic centimeter. At any temperature the atoms of the solid are vibrating. The amplitude of vibration is about $10^{-9}$ cm and the frequency about $10^{13}$/sec.

When the temperature is increased the average distance between atoms increases. This leads to an expansion of the whole solid body as the temperature is increased. The change in *any* linear dimension of the solid, such as its length, width, or thickness, is called a linear expansion. If the

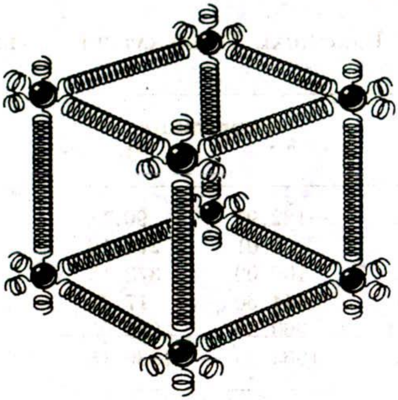


**Fig. 21–5** A solid behaves in many ways as if it is a microscopic "bed-spring" in which the molecules are held together by elastic forces.

length of this linear dimension is $l$, the change in length, arising from a change in temperature $\Delta T$, is $\Delta l$. We find from experiment that, if $\Delta T$ is small enough, this change in length $\Delta l$ is proportional to the temperature change $\Delta T$ and to the original length $l$. Hence, we can write

$$\Delta l = \alpha l \, \Delta T, \qquad (21\text{–}6)$$

where $\alpha$, called the *coefficient of linear expansion*, has different values for different materials. Rewriting this formula we obtain

$$\alpha = \frac{1}{l}\frac{\Delta l}{\Delta T},$$

so that $\alpha$ has the meaning of a fractional change in length per degree temperature change.

Strictly speaking, the value of $\alpha$ depends on the actual temperature and the reference temperature chosen to determine $l$ (see Problem 15). However, its variation is usually negligible compared to the accuracy with which engineering measurements need to be made. We can safely take it as a constant for a given material, independent of the temperature. In Table 21–3 we list the experimental values for the average coefficient of linear expansion of several common solids. For all the substances listed, the change in size consists of an expansion as the temperature rises, for $\bar{\alpha}$ is positive. The order of magnitude of the expansion is about 1 millimeter per meter length per 100 Celsius degrees.*

▶ **Example 2.** A steel metric scale is to be ruled so that the millimeter intervals are accurate to within about $5 \times 10^{-5}$ mm at a certain temperature. What is the maximum temperature variation allowable during the ruling?

From Eq. 21–6,

$$\Delta l = \alpha l \, \Delta T,$$

we have

$$5 \times 10^{-5}\ \text{mm} = (11 \times 10^{-6}/\text{C}^\circ)(1.0\ \text{mm})\ \Delta T$$

in which we have used $\bar{\alpha}$ for steel, taken from Table 21–3. This yields $\Delta T \cong$ 5 C°. The temperature maintained during the ruling process must be maintained when the scale is being used and it must be held constant to within about 5 C°.

Note (see Table 21–3) that if the alloy invar is used instead of steel, then for the

---

* One Celsius degree (1 C°) is a temperature *interval* ($\Delta T$) of one unit measured on a Celsius scale. One degree Celsius (1° C) is a specific temperature reading ($T$) on that scale.

same required tolerance one can permit a temperature variation of about 75 C°; or for the same temperature variation ($\Delta T = 5$ C°) the tolerance achieved would be more than an order of magnitude better. ◀

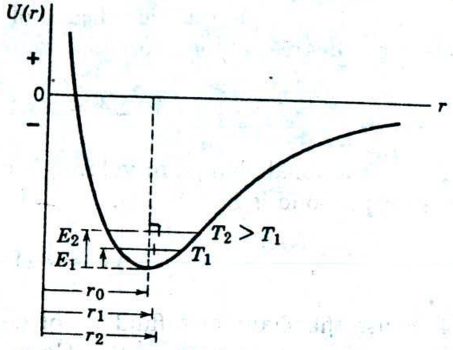
### Table 21–3

SOME VALUES* of $\bar{\alpha}$

| Substance | $\bar{\alpha}$ (per C°) | Substance | $\bar{\alpha}$ (per C°) |
|---|---|---|---|
| Aluminum | $23 \times 10^{-6}$ | Hard rubber | $80 \times 10^{-6}$ |
| Brass | $19 \times 10^{-6}$ | Ice | $51 \times 10^{-6}$ |
| Copper | $17 \times 10^{-6}$ | Invar | $0.7 \times 10^{-6}$ |
| Glass (ordinary) | $9 \times 10^{-6}$ | Lead | $29 \times 10^{-6}$ |
| Glass (pyrex) | $3.2 \times 10^{-6}$ | Steel | $11 \times 10^{-6}$ |

* For the range 0° C to 100° C; except −10° C to 0° C for ice.

On the microscopic level thermal expansion of a solid suggests an increase in the average separation between the atoms in the solid. The potential energy curve for two adjacent atoms in a crystalline solid as a function of their internuclear separation is an asymmetric curve like that of Fig. 21–6. As the atoms move close together, their separation decreasing from the equilibrium value $r_0$, strong repulsive forces come into play and the potential curve rises steeply ($F = -dU/dr$); as the atoms move farther apart, their separation increasing from the equilibrium value, somewhat weaker attractive forces take over and the potential curve rises more slowly. At a given vibrational energy the separation of the atoms will



ig. 21–6 Potential energy curve or two adjacent atoms in a crystal-ine solid as a function of inter-uclear separation. The equilib-ium separation is $r_0$. Because the urve is asymmetric the average eparation ($r_1,r_2$) increases as the emperature ($T_1,T_2$), and hence the ibrational energy ($E_1,E_2$), increases.

hange periodically from a minimum to a maximum value, the average separation eing greater than the equilibrium separation because of the asymmetric nature the potential energy curve. At still higher vibrational energy the average paration will be even greater. The effect is enhanced by the fact that in taking time average of the motion one must allow for the longer time spent at extreme parations (lower vibrational speeds). Because the vibrational energy increases the temperature rises, the average separation between atoms increases with mperature and the solid as a whole expands.

Note that if the potential energy curve were symmetric about the equilibrium separation, then no matter how large the amplitude of the vibration becomes the average separation would correspond to the equilibrium separation. Hence, thermal expansion is a direct consequence of the deviation from symmetry (that is, the asymmetry) of the potential energy curve characteristic of solids.

Some crystalline solids, in certain temperature regions, may contract as the temperature rises. The above analysis remains valid if one assumes that only compressional (i.e. longitudinal) modes of vibration exist or that these modes predominate. However, solids may vibrate in shear-like (i.e. transverse) modes as well and these modes of vibration will allow the solid to contract as the temperature rises, the average separation of the planes of atoms decreasing. For certain types of crystalline structure and in certain temperature regions these transverse modes of vibration may predominate over the longitudinal ones, giving a net negative coefficient of thermal expansion.

It should be emphasized that the microscopic models presented here are over-simplifications of a complex phenomenon which can be treated with greater care with the use of thermodynamics and quantum theory.

For many solids, called *isotropic*, the per cent change in length for a given temperature change is the same for all lines in the solid. The expansion is quite analogous to a photographic enlargement, except that a solid is three-dimensional. Thus, if you have a flat plate with a hole punched in it, $\Delta l/l$ ($= \alpha\Delta T$) for a given $\Delta T$ is the same for the length, the thickness, the face diagonal, the body diagonal, and the hole diameter. Every line, whether straight or curved, lengthens in the ratio $\alpha$ per degree temperature rise. If you scratch your name on the plate, the line representing your name has the same fractional change in length as any other line. The analogy to a photographic enlargement is shown in Fig. 21–7.

With these ideas in mind, the student should be able to show (see Problems 16 and 17) that to a high degree of accuracy the fractional change in area $A$ per degree temperature change for an isotropic solid is $2\alpha$, that is,

$$\Delta A = 2\alpha A\, \Delta T,$$

and the fractional change in volume $V$ per degree temperature change for an isotropic solid is $3\alpha$, that is,

$$\Delta V = 3\alpha V\, \Delta T.$$

Because the shape of a fluid is not definite, only the change in volume with temperature is significant. Gases respond strongly to temperature or pressure changes, whereas the change in volume of liquids with changes in temperature or pressure is very small. If we let $\beta$ represent the coefficient of volume expansion for a liquid, that is,

$$\beta = \frac{1}{V}\frac{\Delta V}{\Delta T},$$

we find that $\beta$ is relatively independent of the temperature. Liquids typi-

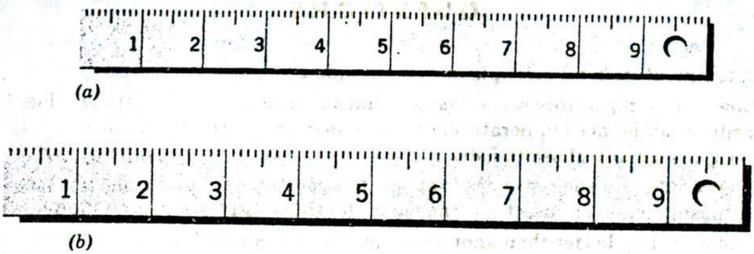


**Fig. 21–7** The same steel rule at two different temperatures. On expansion every dimension is increased by the same proportion: the scale, the numbers, the hole, and the thickness are all increased by the same factor. (The expansion shown, from (a) to (b), is obviously exaggerated, for it would correspond to an imaginary temperature rise of about 100,000 C°!)

cally expand with increasing temperature, their volume expansion being generally about ten times greater than that of solids.

However, the most common liquid, water, does not behave like other liquids. In Fig. 21–8 we show the expansion curve for water. Notice that above 4° C water expands as the temperature rises, although not linearly. As the temperature is lowered from 4 to 0° C, however, water expands instead of contracting. Such an expansion with decreasing temperature is not observed in any other common liquid; it is observed in rubberlike substances and in certain crystalline solids over limited temperature intervals. The density of water is a maximum at 4° C, where its value* is 1000 kg/meter$^3$ or 1.000 gm/cm$^3$. At all other temperatures its density is less. This behavior of water is the reason why lakes freeze first at their upper surface.

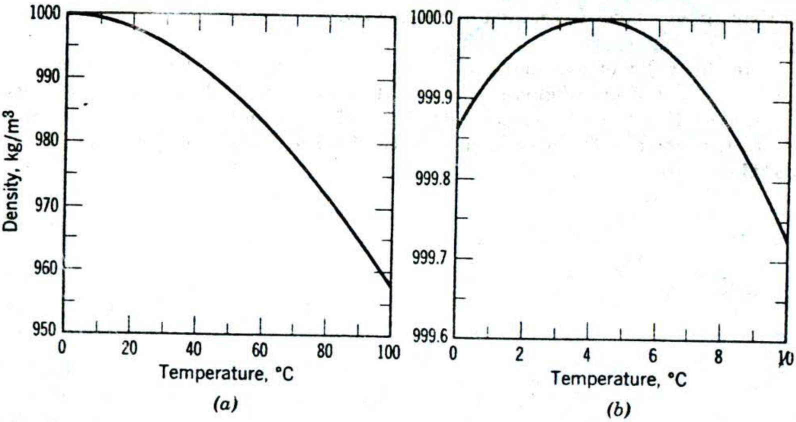


**Fig. 21–8** (a) The variation with temperature of density of water under atmospheric pressure. (b) The variation between 0 and 10°C in more detail.

## QUESTIONS

**1.** Is temperature a microscopic or macroscopic concept?

**2.** Does our "temperature sense" have a built-in sense of direction; that is, does hotter necessarily mean higher temperature, or is this just an arbitrary convention? Celsius, by the way, originally chose the steam point as 0° C and the ice point as 100° C.

**3.** How would you suggest measuring the temperature of (*a*) the sun, (*b*) the earth's upper atmosphere, (*c*) an insect, (*d*) the moon, (*e*) the ocean floor, and (*f*) liquid helium?

**4.** Is one gas any better than another for purposes of a standard constant-volume gas thermometer? What properties are desirable in a gas for such purposes?

**5.** State some objections to using water-in-glass as a thermometer. Is mercury-in-glass an improvement?

**6.** Can you explain why the column of mercury first descends and then rises when a mercury-in-glass thermometer is put in a flame?

**7.** What are the dimensions of $\alpha$, the coefficient of linear expansion? Does the value of $\alpha$ depend on the unit of length used? When F° are used instead of C° as a unit of temperature change, does the numerical value of $\alpha$ change? If so, how?

**8.** A metal ball can pass through a metal ring. When the ball is heated, however, it gets stuck in the ring. What would happen if the ring, rather than the ball, were heated?

**9.** A bimetallic strip, consisting of two different metal strips riveted together, is used as a control element in the common thermostat. Explain how it works.

**10.** Explain how the period of a pendulum clock can be kept constant with temperature by attaching tubes of mercury to the bottom of the pendulum. (See Problem 13).

**11.** Explain why some rubberlike substances contract with rising temperature. (See Question 21, Chapter 25.)

**12.** Explain why the apparent expansion of a liquid in a bulb does not give the true expansion of the liquid.

**13.** Does the change in volume of a body when its temperature is raised depend on whether the body has cavities inside, other things being equal? Consider a solid sphere and a hollow sphere, for example.

**14.** What difficulties would arise if you defined temperature in terms of the density of water?

**15.** Explain why lakes freeze first at the surface.

---

* It is to this value of *unit* maximum density of water that the relative sizes of the kilogram and meter were originally supposed to correspond. Accurate measurements show, however, that the international standards of mass and length do not correspond exactly to this value. The maximum density of water is actually 999.973 kg/meter$^3$ at 3.98° C.

## PROBLEMS

**1.** If the ideal gas temperature at the steam point is 373.15° K, what is the limiting value of the ratio of the pressures of a gas at the steam point and at the triple point of water when the gas is kept at constant volume?

**2.** Let $p_{tr}$ be the pressure in the bulb of a constant-volume gas thermometer when the bulb is at the triple-point temperature of 273.16° K and $p$ the pressure when the bulb is at room temperature. Given three constant-volume gas thermometers: For No. 1 the gas is oxygen and $p_{tr} = 20$ cm-Hg; for No. 2 the gas is also oxygen but $p_{tr} = 40$ cm-Hg; for No. 3 the gas is hydrogen and $p_{tr} = 30$ cm-Hg. The measured values of $p$ for the three thermometers are $p_1$, $p_2$, and $p_3$. (*a*) An approximate value of the room temperature $T$ can be obtained with each of the thermometers using

$$T_1 = 273.16^\circ\,\mathrm{K}\,\frac{p_1}{20\ \text{cm-Hg}}; \quad T_2 = 273.16^\circ\,\mathrm{K}\,\frac{p_2}{40\ \text{cm-Hg}}; \quad T_3 = 273.16^\circ\,\mathrm{K}\,\frac{p_3}{30\ \text{cm-Hg}}.$$

Mark "true" or "false" each of the following statements: (1) With the method described, all three thermometers will give the same value of $T$. (2) The two oxygen thermometers will agree with each other but not with the hydrogen thermometer. (3) Each of the three will give a different value of $T$. (*b*) In the event that there is disagreement among the three thermometers, explain how you would change the method of using them to cause all three to give the same value of $T$.

**3.** It is an everyday observation that hot and cold objects cool down or warm up to the temperature of their surroundings. If the temperature difference $\Delta T$ between an object and its surroundings is not too great, the rate of cooling or warming is approximately proportional to the temperature difference between the object and its surroundings; that is,

$$\frac{d\,\Delta T}{dt} = -K\,\Delta T,$$

where $K$ is a constant. The minus sign appears because $\Delta T$ *decreases* with time if $\Delta T$ is positive and vice versa. This is known as *Newton's law of cooling*. (*a*) On what factors does $K$ depend? What are its dimensions? (*b*) If at some instant the temperature difference is $\Delta T_0$, show that it is

$$\Delta T = \Delta T_0 e^{-Kt}$$

at a time $t$ later.

**4.** A mercury-in-glass thermometer is placed in boiling water for a few minutes and then removed. The temperature readings at various times after removal are as follows:

| $t$, sec | $T$, °C | $t$, sec | $T$, °C | $t$, sec | $T$, °C | $t$, sec | $T$, °C |
|---|---|---|---|---|---|---|---|
| 0 | 98.4 | 25 | 65.1 | 100 | 50.3 | 700 | 26.5 |
| 5 | 76.1 | 30 | 63.9 | 150 | 43.7 | 1000 | 26.1 |
| 10 | 71.1 | 40 | 61.6 | 200 | 38.8 | 1400 | 26.0 |
| 15 | 67.7 | 50 | 59.4 | 300 | 32.7 | 2000 | 26.0 |
| 20 | 66.4 | 70 | 55.4 | 500 | 27.8 | 3000 | 26.0 |

Plot $K$ as a function of time, assuming Newton's law of cooling to apply (see Problem 3). How constant is the "constant" $K$? What might give rise to the observed variations of $K$ with time?

**5.** At what temperature do the Fahrenheit and Celsius scales give the same reading? The Fahrenheit and the Kelvin scales?

**6.** (*a*) The temperature of the surface of the sun is about 6000° K. Express this on the Fahrenheit scale. (*b*) Express normal human body temperature, 98.6° F, on the Celsius scale. (*c*) Excluding Hawaii and Alaska, the highest recorded temperature in the United States is 134° F at Death Valley, California, and the lowest is −70° F at Rogers Pass, Montana. Express these extremes on the Celsius scale. (*d*) Express the normal boiling point of oxygen, −183° C, on the Fahrenheit scale. (*e*) At what Celsius temperature would you find a room to be uncomfortably warm?

**7.** In the interval between 0 and 660° C, a platinum resistance thermometer of definite specifications is used for interpolating temperatures on the International Practical Temperature Scale. The temperature $t$ is given by a formula for the variation of resistance with temperature:

$$R = R_0(1 + At + Bt^2).$$

$R_0$, $A$, and $B$ are constants determined by measurements at the ice point, the steam point, and the sulphur point. (*a*) If $R$ equals 10.000 ohms at the ice point, 13.946 ohms at the steam point, and 24.817 ohms at the sulphur point, find $R_0$, $A$, and $B$. (*b*) Plot $R$ versus $t$ in the temperature range from 0 to 660° C.

**8.** (*a*) Show that if the lengths of two rods of different solids are inversely proportional to their respective coefficients of linear expansion at some initial temperature, the difference in length between them will be constant at all temperatures. (*b*) What should be the lengths of a steel and a brass rod at 0° C so that at all temperatures their difference in length is 0.30 meter?

**9.** A circular hole in an aluminum plate is 1.000 in. in diameter at 0° C. What is its diameter when the temperature of the plate is raised to 100° C?

**10.** The Pyrex glass mirror in the telescope at Palomar Observatory has a diameter of 200 in. The temperature ranges from −10 to 50° C on Mount Palomar. Determine the maximum change in the diameter of the mirror.

**11.** A clock pendulum made of invar has a period of 0.500 sec at 20° C. If the clock is used in a climate where the temperature averages 30° C, what correction (approximately) is necessary at the end of 30 days to the time given by the clock?

**12.** The distance between the towers of the main span of the Golden Gate Bridge at San Francisco is 4200 ft. The sag of the cable halfway between the towers at 50° F is 470 ft. Take $\bar{\alpha} = 6.5 \times 10^{-6}/\text{F}^\circ$ for the cable and compute the change in length of the cable and the change in sag for a temperature change from −20 to 110°F. Assume no bending or separation of the towers and a parabolic shape for the cable.

**13.** A glass tube nearly filled with mercury is attached to the bottom of an iron pendulum rod 100 cm long. How high must the mercury be in the glass tube so that the center of mass of this pendulum will not rise or fall with changes in temperature?

**14.** A steel rod is 3.000 cm in diameter at 25° C. A brass ring has an interior diameter of 2.992 cm at 25° C. At what common temperature will the ring just slide onto the rod?

**15.** Show that if $\alpha$ is treated as a variable, dependent on the temperature $T$, then

$$L = L_0\left[1 + \int_{T_0}^{T} \alpha(T)\, dT\right]$$

where $L_0$ is the length at a reference temperature $T_0$.

**16.** The area $A$ of a rectangular plate is $ab$. Its coefficient of linear expansion is $\alpha$. After a temperature rise $\Delta T$, side $a$ is longer by $\Delta a$ and side $b$ is longer by $\Delta b$. Show that if we neglect the small area $\Delta a \cdot \Delta b$, shown cross-hatched and greatly exaggerated in size in Fig. 21-9, then $\Delta A = 2\alpha A\,\Delta T$.

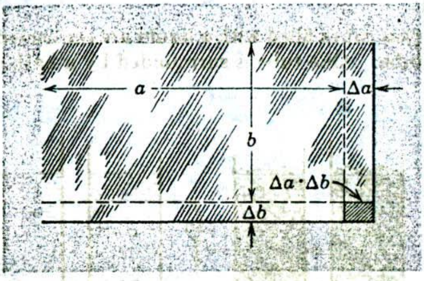


Fig. 21–9

**17.** Prove that, if we neglect extremely small quantities, the change in volume of a solid on expansion through a temperature rise $\Delta T$ is given by $\Delta V = 3\alpha V\,\Delta T$ where $\alpha$ is the coefficient of linear expansion.

**18.** When the temperature of a "copper" penny is raised by 100° C, its diameter increases by 0.18%. To two significant figures give the per cent increase in the (*a*) area of a face, (*b*) thickness, (*c*) volume, and (*d*) mass of the penny. (*e*) What is the coefficient of linear expansion?

**19.** Find the change in volume of an aluminum sphere of 10.0-cm radius when it is heated from 0.00 to 100° C.

**20.** Consider a mercury-in-glass thermometer. Assume that the cross-section of the capillary is constant at $A_0$, and that $V_0$ is the volume of the bulb of mercury at 0.00° C. If the mercury just fills the bulb at 0.00° C, show that the length of the mercury column in the capillary at a temperature $t$° C is

$$l = \frac{V_0}{A_0}(\beta - 3\alpha)t,$$

that is, proportional to the temperature, where $\beta$ is the volume coefficient of expansion of mercury and $\alpha$ is the linear coefficient of expansion of glass.

**21.** Density is mass per unit volume. If the volume $V$ is temperature dependent, so is the density $\rho$. Show that the change in density $\Delta\rho$ with change in temperature $\Delta T$ is given by

$$\Delta\rho = -\beta\rho\,\Delta T$$

where $\beta$ is the volume coefficient of expansion. Explain the minus sign.

**22.** (*a*) Prove that the change in rotational inertia $I$ with temperature of a solid object is given by $\Delta I = 2\alpha I\,\Delta T$. (*b*) Prove that the change in period $t$ of a physical pendulum with temperature is given by $\Delta t = \frac{1}{2}\alpha t\,\Delta T$.

**23.** Consider a uniform solid brass cylinder of mass $M = 0.50$ kg and radius $R = 0.030$ meter. The cylinder is placed in frictionless bearings and set to rotate about its cylinder axis with an angular velocity $\omega = 60$ radians/sec. (*a*) What is the angular momentum of the cylinder and how much work is required to reach this rate of rotation, starting from rest? (*b*) After the cylinder has reached the state of rotation just described we heat it, without mechanical contact, from room temperature (20° C) to 100° C. Take the mean coefficient of linear expansion of brass to be $\bar{\alpha} = 2.0 \times$

$10^{-5}$/C°. Find the fractional changes, if any, in the angular velocity, the angular momentum, and the kinetic energy of rotation of the cylinder. Explain.

**24.** Show that when the temperature of a liquid in a barometer changes by $\Delta T$, and the pressure is constant, the height $h$ changes by $\Delta h = \beta h \, \Delta T$ where $\beta$ is the coefficient of volume expansion.

**25.** Two vertical glass tubes filled with a liquid are connected at their lower ends by a horizontal capillary tube. One tube is surrounded by a bath containing ice and water in equilibrium (0.0° C), the other by a hot-water bath ($t$). The difference in height of the liquids in the two columns is $\Delta h$, and $h_0$ is the height of the column at 0.0° C. Show how this apparatus (Fig. 21–10), first used in 1816 by Dulong and Petit, can be used to measure the true coefficient of volume expansion $\beta$ of a liquid (rather than the differential expansion between glass and liquid). Determine $\beta$ if $t$ = 16.0° C, $h_0$ = 126 cm, and $\Delta h$ = 1.50 cm.

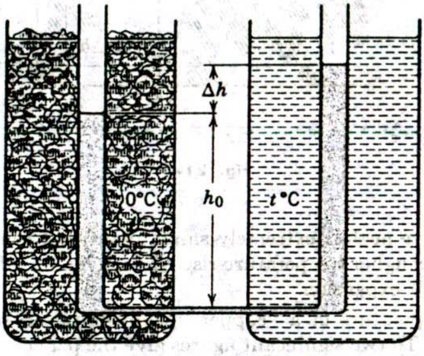


**Fig. 21–10**

# Heat and the First Law of Thermodynamics

## CHAPTER 22

### 22–1 Heat, a Form of Energy

When two systems at different temperatures are placed together, the final temperature reached by both systems is somewhere between the two starting temperatures. This is a common observation. Man has long sought for a deeper understanding of such phenomena. Up to the beginning of the nineteenth century, they were explained by postulating that a material substance, *caloric*, existed in every body. It was believed that a body at high temperature contained more caloric than one at a low temperature. When the two bodies were put together, the body rich in caloric lost some to the other until both bodies reached the same temperature. The caloric theory was able to describe many processes, such as heat conduction or the mixing of substances in a calorimeter, in a satisfactory way. However, the concept of heat as a *substance*, whose total amount remained constant, eventually could not stand the test of experiment. Nevertheless, we still describe many common temperature changes as the transfer of "something" from one body at a higher temperature to one at the lower, and this "something" we call heat. A useful but nonoperational definition, is: ***heat is that which is transferred between a system and its surroundings as a result of temperature differences only.***

Eventually it became generally understood that heat is a form of energy rather than a substance. The first conclusive evidence that heat could not be a substance was given by Benjamin Thompson (1753–1814), an American who later became Count Rumford of Bavaria. In a paper read before the Royal Society* in 1798 he wrote:

> I . . . am persuaded, that a habit of keeping the eyes open to everything that is going on in the ordinary course of the business of life has oftener led, as it were by accident, or in the playful excursions of the imagination . . . to useful doubts and sensible schemes for investigation and improvement, than all the more intense meditations of philosophers, in the hours expressly set apart for study. It was by accident that I was led to make the Experiments of which I am about to give an account.

Rumford made his discovery while supervising the boring of cannon for the Bavarian government. To prevent overheating, the bore of the cannon was kept full of water. The water was replenished as it boiled away during the boring process. It was accepted that caloric had to be supplied to water to boil it. The continuous production of caloric was explained by assuming that when a substance was more finely subdivided, as in boring, its capacity for retaining caloric became smaller, and that the caloric released in this way was what caused the water to boil. Rumford observed in specific experiments, however, that the water boiled away even when his boring tools became so dull that they were no longer cutting or subdividing matter.

He wrote after ruling out by experiment all possible caloric interpretations,

> . . . in reasoning on this subject, we must not forget to consider that most remarkable circumstance, that the source of Heat generated by friction, in these Experiments, appeared evidently to be *inexhaustible* . . . it appears to me to be extremely difficult, if not quite impossible, to form any distinct idea of any thing capable of being excited and communicated in the manner the Heat was excited and communicated in these Experiments, except it be *MOTION*.

Here we have the germ of the idea that the mechanical work expended in the boring process was responsible for the creation of heat. The idea was not clearly put until much later, by others. Instead of the continuous disappearance of mechanical energy and the continuous creation of heat, neither obeying any conservation principle, the whole process is now viewed as a transformation of energy from one form to another, the total energy being conserved.

Although the concept of energy and its conservation seems self-evident today, it was a novel idea as late as the 1850's and had eluded such men as Galileo and Newton. Throughout the subsequent history of physics this conservation idea led men to new discoveries. Its early history was remarkable in many ways. Several thinkers arrived at this great concept at about the same time; at first, all of them either met with a cold reception or were ignored. The principle of the conservation of energy

---

* Rumford, an American, was founder of the Royal Institution in London. On the other hand, the Smithsonian Institution in Washington owes its origin to an Englishman.

was established independently by Julius von Mayer (1814–1878) in Germany, James Joule (1818–1889) in England, Hermann von Helmholtz (1821–1894) in Germany, and L. A. Colding (1815–1888) in Denmark.*

It was Joule who showed by experiment that, when a given quantity of mechanical energy is converted to heat, the same quantity of heat is always developed. Thus, the equivalence of heat and mechanical work as two forms of energy was definitely established.

Helmholtz first expressed clearly the idea that not only heat and mechanical energy but all forms of energy are equivalent, and that a given amount of one form cannot disappear without an equal amount appearing in some of the other forms.

## 22–2 Quantity of Heat and Specific Heat

The unit of heat $Q$ is defined quantitatively in terms of a specified change produced in a body during a specified process. Thus, if the temperature of one kilogram of water is raised from 14.5 to 15.5° C by heating, we say that one *kilocalorie* (kcal) of heat has been added to the system. The *calorie* ($= 10^{-3}$ kcal) is also used as a heat unit. (Incidentally, the "calorie" used to measure the energy content of foods is actually a kilocalorie.) In the engineering system the unit of heat is the *British thermal unit* (Btu), which is defined as the heat necessary to raise the temperature of one pound of water from 63 to 64° F.

The reference temperatures are stated because, near room temperature, there is a slight variation in the heat needed for a one-degree temperature rise with the temperature interval chosen. We will neglect this variation for most practical purposes. The heat units are related as follows:

$$1.000 \text{ kcal} = 1000 \text{ cal} = 3.968 \text{ Btu}$$

Substances differ from one another in the quantity of heat needed to produce a given rise of temperature in a given mass. The ratio of the heat $\Delta Q$ supplied to a body to its corresponding temperature rise $\Delta T$ is called the *heat capacity* $C$ of the body; that is,

$$C = \text{heat capacity} = \frac{\Delta Q}{\Delta T}.$$

The word "capacity" may be misleading because it suggests the essentially meaningless statement "the amount of heat a body can hold," whereas what is meant is simply the heat added per unit temperature rise.

---

* From the posthumous publication of *Reflections* (1872) of the French engineer Sadi Carnot (1796–1832), it is clear that he arrived at the conservation of energy principle before all the others. It will give the student some food for thought to realize that of the five men who were the first to understand the conservation of energy principle, all were young and all were professionally outside the field of physics at the time of their contributions. Mayer was a physician, age 28; Helmholtz, a physiologist, age 32; Colding, an engineer, age 27; Joule, an industrialist, age 25; and Carnot, an engineer, age 34. Rumford was an old man, age 45, by comparison.

The heat capacity per unit mass of a body, called *specific heat*, is characteristic of the material of which the body is composed:

$$c = \frac{\text{heat capacity}}{\text{mass}} = \frac{\Delta Q}{m\,\Delta T}. \tag{22–1}$$

We properly speak, on the one hand, of the heat capacity of a penny but, on the other, of the specific heat of copper.

Neither the heat capacity of a body nor the specific heat of a material is constant but depends on the location of the temperature interval. The previous equations give only average values for these quantities in the temperature range of $\Delta T$. The specific heat $c$ of a material *at any temperature* is defined by

$$c = \frac{dQ}{m\,dT}. \tag{22–2}$$

Hence, the heat that must be given to a body of mass $m$, whose material has a specific heat capacity $c$, to increase its temperature from $T_i$ to $T_f$, is

$$Q = m \int_{T_i}^{T_f} c\,dT \tag{22–3}$$

where $c$ is a function of the temperature. At ordinary temperatures and over ordinary temperature intervals, specific heats can be considered to be constants. Figure 22–1 shows the variation in the specific heat of water with temperature. Information of this sort is obtained by using an electrical heating coil to supply heat at a rate that can be accurately determined. We see from the graph that the specific heat of water varies less than 1% from its value of 1.000 cal/gm C° at 15° C.

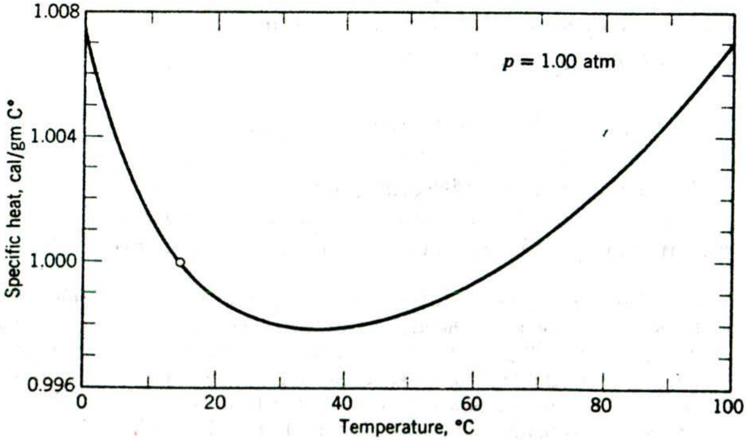


**Fig. 22–1** The variation with temperature of the specific heat of water at a pressure of 1.00 atm. The circle, located at 15° C, suggests the definition of the calorie.

Equations 22-1 and 22-2 do not define specific heat uniquely. We must also specify the conditions under which the heat $\Delta Q$ is added to the specimen. We have implied that the condition is that the specimen remain at normal (constant) atmospheric pressure while we add the heat. This is a common condition, but there are many other possibilities, each leading, in general, to a different value for $c$. To obtain a unique value for $c$ we must specify the conditions, such as specific heat at constant pressure $c_p$, specific heat at constant volume $c_v$, etc.

Table 22-1 (second column) shows the specific heats at constant pressure of some solid elements; we will discuss the specific heats of gases later. The student should realize from the way the calorie and the Btu are defined that 1 cal/gm C° = 1 kcal/kg C° = 1 Btu/lb F°, exactly. Note that the specific heat of water, equal to 1.00 cal/gm C°, is large compared to that of most substances.

**Table 22-1**

VALUES FOR $c_p$ FOR SOME SOLIDS
(At room temperature and for $p = 1.0$ atm)

| Substance | Specific heat, cal/gm C° | Molecular weight gm/mole | Molar heat capacity cal/mole C° |
|---|---|---|---|
| Aluminum | 0.215 | 27.0 | 5.82 |
| Carbon | 0.121 | 12.0 | 1.46 |
| Copper | 0.0923 | 63.5 | 5.85 |
| Lead | 0.0305 | 207 | 6.32 |
| Silver | 0.0564 | 108 | 6.09 |
| Tungsten | 0.0321 | 184 | 5.92 |

▶ **Example 1.** A 75-gm block of copper, taken from a furnace, is dropped into a 300-gm glass beaker containing 200-gm of water. The temperature of the water rises from 12 to 27° C. What was the temperature of the furnace?

This is an example of two systems originally at different temperatures reaching thermal equilibrium after contact. No mechanical energy is involved, only heat exchange. Hence,

heat lost by copper = heat gained by (beaker + water),

$$m_C c_C (T_C - T_e) = (m_G c_G + m_W c_W)(T_e - T_W).$$

The subscript $C$ stands for copper, $G$ for glass, and $W$ for water. The initial copper temperature is $T_C$, the initial beaker water temperature is $T_W$, and $T_e$ is the final equilibrium temperature. Substituting the given values, with $c_C = 0.093$ cal/gm C°, $c_G = 0.12$ cal/gm C°, and $c_W = 1.0$ cal/gm C°, we obtain

$$(75 \text{ gm})(0.093 \text{ cal/gm C}^\circ)(T_C - 27^\circ \text{ C}) = [(300 \text{ gm})(0.12 \text{ cal/gm C}^\circ) + (200 \text{ gm})(1.0 \text{ cal/gm C}^\circ)](27^\circ \text{ C} - 12^\circ \text{ C})$$

or, solving for $T_C$, $$T_C = 530^\circ \text{ C}$$

What approximations, both experimental and theoretical, were used implicitly to arrive at this answer? ◀

## 22–3 Molar Heat Capacities of Solids

From the second column of Table 22–1 we conclude that the specific heats of solids vary widely from one material to another. However quite a different story emerges if we compare samples of materials that contain the same number of molecules rather than samples that have the same mass. We can do this by expressing specific heats (called when so expressed *molar heat capacities*) in cal/mole C° rather than in cal/gm C°.* In 1819 Dulong and Petit pointed out that the molar heat capacities of all substances, with few exceptions (see carbon in Table 22–1), have values close to 6 cal/mole C°. The molar heat capacity, listed in the fourth column of Table 22–1, is found by multiplying the specific heat (second column) by the molecular weight (third column). We see that the amount of heat required *per molecule* to raise the temperature of a solid by a given amount seems to be about the same for almost all materials. This is striking evidence for the molecular theory of matter.

Actually molar heat capacities vary with temperature, approaching zero as $T \rightarrow 0°$ K and approaching the Dulong-Petit value as $T \rightarrow \infty$. Since the number of molecules rather than the kind of molecule seems to be important in determining the heat required to increase the temperature of a body by a given amount, we are led to expect that the molar heat capacities of different substances will vary with temperature in much the same way. Figure 22–2 shows that, indeed, the molar heat capacities of various substances can be made to fall on the same curve by a simple, empirical adjustment in the temperature scale. The horizontal scale in Fig. 22–2 is the dimensionless ratio $T/T_D$, where $T$ is the Kelvin temperature and $T_D$ is a characteristic temperature, called the *Debye temperature,* that has a particular constant value for each material. For lead, $T_D$ has the empirical value of 88° K and for carbon, $T_D = 1860°$ K. From these data the student can show that a scale value of $T/T_D = 0.600$ corresponds to $T = 53°$ K for lead but to $T = 1120°$ K for carbon. Alternatively, room temperature ($\sim$300° K) corresponds to $T/T_D = 3.4$ for lead and to $T/T_D = 0.16$ for carbon. Thus we see from Fig. 22–2 that in the early days, when only room temperature specific heats were available, lead would conform to the Dulong and Petit rule but carbon would seem to be an exception.

The straight line $I$ in Fig. 22–2 is the Dulong and Petit value of 1819; it agrees with experiment at high temperature but fails at low temperatures. It corresponds to the assumption that every atom in a solid vibrates independently like a classical oscillator. Curve II is due to Debye (1912). In the Debye theory, a characteristic temperature $T_D$, which is directly related to a vibrational frequency characteristic of the material, can be obtained independent of specific heat experi-

---

* A *mole* of any substance is that mass of the substance that contains a specified number of molecules, namely, $6.02252 \times 10^{23}$, called Avogadro's number. This number is the result of the defining relation that one mole of carbon (actually, of the isotope $C^{12}$) shall have a mass of 12 gm, exactly. The *molecular weight* $M$ of a substance is a dimensionless quantity expressing the number of grams per mole of that substance. Thus the molecular weight of oxygen is 32.0 gm/mole. Although the mole is a unit of mass, we cannot translate it into, say, grams, until we know the chemical composition of the substance; for this reason we find it convenient to use a special symbol ($\mu$) for masses expressed in moles.

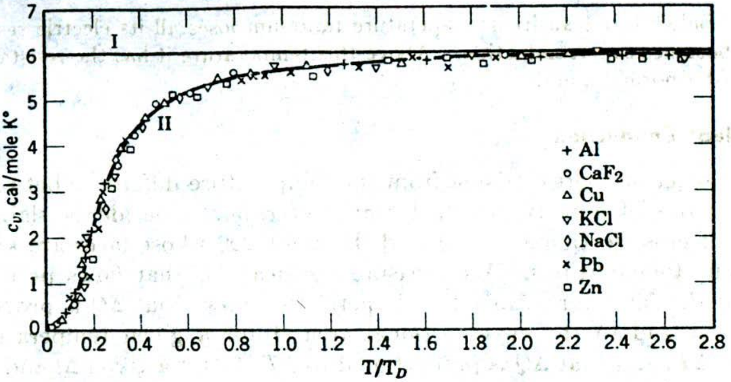
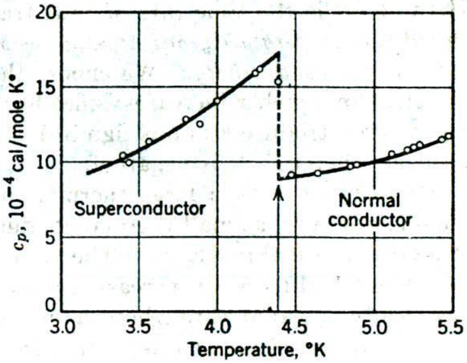


**Fig. 22-2** The molar heat capacities ($c_v$) showing a few selected points only. Line I represents the Dulong and Petit rule and curve II represents a theory due to Debye.

ments. One then uses quantum principles to analyze the coupled vibrations of the atoms in a solid and obtains a specific heat formula which, in terms of the dimensionless ratio $T/T_D$, is the same for all substances. The excellent agreement of this formula (curve II) with experiment is a triumph of quantum physics.*

The materials displayed in Fig. 22-2 are "normal" in that they do not melt, boil, change their crystal structure, etc., in the temperature range indicated. Specific heat measurements, which tell us how a solid absorbs energy as its temperature is raised, are a sensitive probe to detect such molecular, atomic, or electronic rearrangements. Figure 22-3, for example, shows the specific heat of tantalum near

**Fig. 22-3** The specific heat of tantalum near its superconducting transition temperature.



---

* The data reported in Fig. 22-2 are values of $c_v$ but those in Table 22-1 are $c_p$. The former is easier to calculate theoretically because the thermal expansion need not be taken into account, but (for solids) the latter is much easier to measure. The two are related by the simple thermodynamic formula

$$c_p = c_v + T\beta^2/\kappa\rho$$

in which $\beta$ is the thermal coefficient of volume expansion, $\kappa$ ($= -\Delta V/V\Delta p$) is the (isothermal) compressibility, and $\rho$ is the density. At room temperature the difference between $c_p$ and $c_v$ for typical solids is about 5%.

4.39° K. Below this transition temperature tantalum loses all its electric resistance—it becomes superconducting. Above this temperature it has the resistance expected of a normal metal.

## 22–4 Heat Conduction

The transfer of energy arising from the temperature difference between adjacent parts of a body is called *heat conduction*. Consider a slab of material of cross-sectional area $A$ and thickness $\Delta x$, whose faces are kept at different temperatures. We measure the heat $\Delta Q$ that flows perpendicular to the faces for a time $\Delta t$. Experiment shows that $\Delta Q$ is proportional to $\Delta t$ and to the cross-sectional area $A$ for a given temperature difference $\Delta T$, and that $\Delta Q$ is proportional to $\Delta T/\Delta x$ for a given $\Delta t$ and $A$, providing both $\Delta T$ and $\Delta x$ are small. That is,

$$\frac{\Delta Q}{\Delta t} \propto A\,\frac{\Delta T}{\Delta x} \qquad \text{approximately.}$$

In the limit of a slab of infinitesimal thickness $dx$, across which there is a temperature difference $dT$, we obtain the fundamental law of heat conduction

$$\frac{dQ}{dt} = -kA\,\frac{dT}{dx}. \tag{22–4}$$

Here $dQ/dt$ is the time rate of heat transfer across the area $A$, $dT/dx$ is called the *temperature gradient*, and $k$ is a constant of proportionality called the *thermal conductivity*. We choose the direction of heat flow to be the direction in which $x$ increases; since heat flows in the direction of decreasing $T$, we introduce a minus sign in Eq. 22–4 (that is, we wish $dQ/dt$ to be positive when $dT/dx$ is negative).

A substance with a large thermal conductivity $k$ is a good heat conductor; one with a small thermal conductivity $k$ is a poor heat conductor, or a good thermal insulator. The value of $k$ depends on the temperature, increasing slightly with increasing temperature, but $k$ can be taken to be practically constant throughout a substance if the temperature difference between its parts is not too great. In Table 22–2 we list values of $k$ for various substances; we see that metals as a group are better heat conductors than nonmetals, and that gases are poor heat conductors.

Let us apply Eq. 22–4 to a rod of length $L$ and constant cross-sectional area $A$ in which a steady state has been reached (Fig. 22–4). In a steady state the temperature at each point is constant in time. Hence, $dQ/dt$ is the same at all cross-sections. (Why?) But $dQ/dt = -kA(dT/dx)$, so that, for a constant $k$ and $A$, the temperature gradient $dT/dx$ is the same at all cross-sections. Hence, $T$ decreases linearly along the rod so that

### Table 22-2

THERMAL CONDUCTIVITIES, KCAL/SEC METER C°
(Gases at 0° C; others at about room temperature)

| | |
|---|---|
| Metals | |
| Aluminum | $4.9 \times 10^{-2}$ |
| Brass | $2.6 \times 10^{-2}$ |
| Copper | $9.2 \times 10^{-2}$ |
| Lead | $8.3 \times 10^{-3}$ |
| Silver | $9.9 \times 10^{-2}$ |
| Steel | $1.1 \times 10^{-2}$ |
| Gases | |
| Air | $5.7 \times 10^{-6}$ |
| Hydrogen | $3.3 \times 10^{-5}$ |
| Oxygen | $5.6 \times 10^{-6}$ |
| Others | |
| Asbestos | $2 \times 10^{-5}$ |
| Concrete | $2 \times 10^{-4}$ |
| Cork | $4 \times 10^{-5}$ |
| Glass | $2 \times 10^{-4}$ |
| Ice | $4 \times 10^{-4}$ |
| Wood | $2 \times 10^{-5}$ |

$-dT/dx = (T_2 - T_1)/L$. Therefore, the heat $\Delta Q$ transferred in time $\Delta t$ is

$$\frac{\Delta Q}{\Delta t} = kA\,\frac{T_2 - T_1}{L}. \tag{22-5}$$

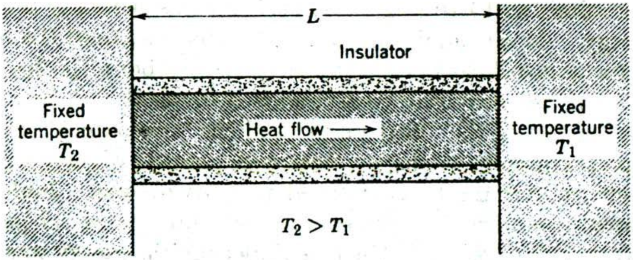


**Fig. 22-4** Conduction of heat through an insulated conducting bar.

The phenomenon of heat conduction also shows that the concepts of heat and temperature are distinctly different. Different rods, having the same temperature difference between their ends, may transfer entirely different quantities of heat in the same time.

▶ **Example 2.** Consider a compound slab, consisting of two materials having different thicknesses, $L_1$ and $L_2$, and different thermal conductivities, $k_1$ and $k_2$. If the temperatures of the outer surfaces are $T_2$ and $T_1$, find the rate of heat transfer through the compound slab (Fig. 22-5) in a steady state.

Let $T_x$ be the temperature at the interface between the two materials. Then

$$\frac{\Delta Q_2}{\Delta t} = \frac{k_2 A (T_2 - T_x)}{L_2}$$

and

$$\frac{\Delta Q_1}{\Delta t} = \frac{k_1 A (T_x - T_1)}{L_1}.$$

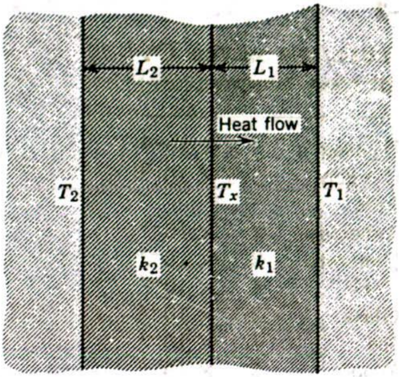


**Fig. 22–5** Example 2. Conduction of heat through two layers of matter with different thermal conductivities.

In a steady state $\Delta Q_1/\Delta t = \Delta Q_2/\Delta t$, so that

$$\frac{k_2 A(T_2 - T_x)}{L_2} = \frac{k_1 A(T_x - T_1)}{L_1}.$$

Let $\Delta Q/\Delta t$ be the rate of heat transfer (the same for all sections). Then, solving for $T_x$ and substituting into either of these equations, we obtain

$$\frac{\Delta Q}{\Delta t} = \frac{A(T_2 - T_1)}{(L_1/k_1) + (L_2/k_2)}.$$

The extension to any number of sections in series is obviously

$$\frac{\Delta Q}{\Delta t} = \frac{A(T_2 - T_1)}{\Sigma(L_i/k_i)}.$$ ◀

## 22–5 The Mechanical Equivalent of Heat

If heat is just another form of energy, any energy unit could be a heat unit. The calorie and Btu originated before it was generally accepted that heat is energy. It was Joule who first carefully measured the mechanical energy equivalent of heat energy, that is, the number of joules equivalent to 1 calorie, or the number of foot-pounds equivalent to 1 Btu.

The relative size of the "heat units" and the "mechanical units" can be found from experiments in which a measured quantity of mechanical energy is completely converted into a measured quantity of heat. Joule originally used an apparatus in which falling weights rotated a set of paddles in a water container (Fig. 22–6). The loss of mechanical energy was computed

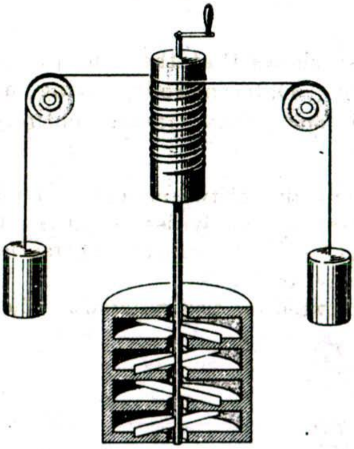


**Fig. 22–6** Joule's arrangement for measuring the mechanical equivalent of heat. The falling weights turn paddles which stir the water in the container, thus raising its temperature.

from a knowledge of the weights and the heights through which they fell and the gain in heat energy by determining the equivalent mass of water and its rise in temperature. Joule wanted to show that the same amount of heat energy would be obtained from a given expenditure of work regardless of the method used to produce the work. He produced heat by stirring mercury, by rubbing together iron rings in a mercury bath, by converting electrical energy into heat in a wire immersed in water, and in other ways. Always the constant of proportionality between heat produced and work performed agreed within his experimental error of 5%. Joule did not have at his disposal the accurately standardized thermometers of today, nor could he make such reliable corrections for heat losses from the system as are possible now. His pioneer experiments are noteworthy not only for the skill and ingenuity he showed but also for the influence they had in convincing scientists everywhere of the correctness of the concept that heat is a form of energy.

The accepted results* (see Appendix A for more precise values) are

$$\begin{aligned}1 \text{ kcal} &= 1000 \text{ cal} = 4186 \text{ joules}\\ 1 \text{ Btu} &= 252.0 \text{ cal} = 777.9 \text{ ft lb};\end{aligned}$$

that is, 4186 joules of mechanical energy, when converted to heat, will raise the temperature of 1 kg of water from 14.5 to 15.5° C.

In modern calorimetry heat is almost always measured in terms of the electrical energy transferred to a water bath by passing a current through a resistor that is immersed in the bath; it is rarely measured by observing the rise in temperature of a water bath. Thus the logical practical unit of heat is the joule (1 joule = 1 watt-sec) and this was indeed adopted as the accepted international unit for heat by the Ninth General Conference on Weights and Measures (1948). Indeed, in modern laboratory practice the calorie (or kilocalorie) is not much used or needed. It is, however, deeply embedded in the literature of science. To permit the continued use of this familiar unit—but to recognize the practical importance of the joule—a new kilocalorie, the *thermochemical kilocalorie,* is often defined:

$$1 \text{ kilocalorie (thermochemical)} = 4184.0 \text{ joules (exactly)}.$$

In ordinary laboratory practice this kilocalorie does not differ significantly from that defined earlier.

## 22-6 Heat and Work

We have seen that *heat is energy that flows from one body to another because of a temperature difference between them.* The idea that heat is something *in* a body, as the caloric theory assumed, contradicts many experimental facts. It is only as it flows, because of a temperature difference, that the energy

---

* Henry A. Rowland, in 1879, carried out a painstaking determination of the mechanical equivalent of heat which, to this day, remains a model of careful experimentation. His result differs from the accepted value today by only 1 part in 2000. Rowland graduated from Rensselaer Polytechnic Institute in 1870 and in 1876 became the first Professor of Physics at the then newly established Johns Hopkins University, where he conducted this experiment. See "The Education of an American Scientist, Henry A. Rowland" by Samuel Rezneck, *American Journal of Physics,* February 1960.

is called heat energy. If heat were a substance, or a definite kind of energy that kept its identity while contained in a system, it would not be possible to remove heat indefinitely from a system which does not change. Yet Rumford showed that this was possible. In fact, by continually performing mechanical work in Joule's apparatus, we can obtain an indefinite amount of heat out of the water, by connecting it to a cooler system, for example, without changing the condition of the water.

In the same way work is not something of which a system contains a definite amount. We can put an indefinite amount of work into a system without changing its condition, as Joule's apparatus again illustrates. Work, like heat, involves a transfer of energy. In mechanics, work is involved in energy transfers in which temperature played no role. If heat energy is transmitted by temperature differences, we can distinguish heat and work by defining *work as energy that is transmitted from one system to another in such a way that a difference of temperature is not directly involved.* This definition is consistent with our previous use of the term. That is, in the expression $dW = F\,dx$, the force $F$ can arise from electrical, magnetic, gravitational, and other sources. The term work includes all these energy transfer processes, but it specifically excludes energy transfer arising from temperature differences.

Consider another simple example, that of rubbing two surfaces together. There is no limit to the amount of heat that can be removed from this system or to the amount of work that can be put into it, so that there is no definite meaning to phrases such as "the heat in the system" or "the work in the system." The quantities $Q$ and $W$ are not characteristic of the (equilibrium) *state* of the system but rather of the *thermodynamic process* by which the system moves from one equilibrium state to another, by interacting with its environment. It is only during such a process that we can give meaning to heat and work; we can then identify $Q$ with the heat transferred to or from the system and $W$ with the work done on or by the system. The study of such processes and of the changes in energy involved in the performance of work and the flow of heat is the subject matter of *thermodynamics.*

In Fig. 22–7 we consider a general thermodynamic process. We must first state definitely what the system is and what the environment is. In the figure we draw a closed surface surrounding the system to define it. In (*a*) the system is in its *initial state,* in equilibrium with the environment external to it. In (*b*) the system interacts with its environment through some specific *thermodynamic process.* During this process, energy in the form of heat and/or work may go into or out of the system. Arrows representing the flow of $Q$ or $W$ must pierce the surface enclosing the system. In (*c*) the system has reached its *final state,* again in equilibrium with the environment external to it.

Figure 22–8 shows a falling weight which turns a generator, which in turn sends an electric current through a resistor immersed in a water container. Let us choose the system to be the generator and the attached electric circuit, the water, and its container Then the environment is the weight and the earth, which pulls on the weight. The process consists of letting the weight fall a distance $h$ in the earth's gravitational field. During this process the environment (by means of the cord) does work $W$ on the system. There are no temperature differences between the system and its environment and hence $Q = 0$ for this process.

Our choice of a system in thermodynamic problems is arbitrary. Let us now choose the system to be only the water and its container in Fig. 22–8. The environment now is the generator and attached circuit as well as the weight and the earth. For this choice of system there now *is* a temperature difference between the environment (resistor) and the system (water), and heat $Q$ will flow into the system during the process. No forces act through the system boundary to product displacements, however, and hence $W = 0$ for this process. This example shows that we must first state definitely what the system is and what the environment is before we can decide whether the change in the state of the system is due to the flow of heat or to the performance of work or both. There will be a transfer of heat between system and environment only when a temperature difference exists across the system boundary; if no temperature difference exists, the energy transfer involves work.

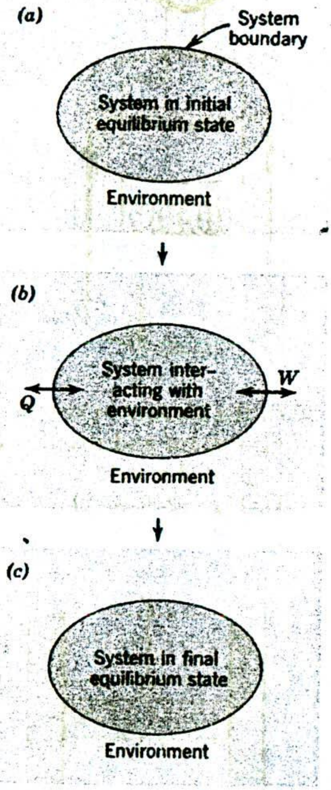


**Fig. 22–7** (*a*) A system in an initial state, in equilibrium with its surroundings. (*b*) A thermodynamic process during which the system may exchange heat $Q$ or work $W$ with its environment. (*c*) A final equilibrium state reached as the result of the process.

Let us now compute $Q$ and $W$ for a specific thermodynamic process. Consider a gas in a cylindrical container with a movable piston. Let the gas be the system. Initially it is in equilibrium with the environment external to it (which is the heat reservoir and the piston, shown in Fig. 22–9) and has a pressure $p_i$ and a volume $V_i$. We can think of the containing walls as the system boundary. Heat can flow into the system or out of it through the bottom of the cylinder and work can be done on the system or by the system by compressing or expanding the gas, respectively,

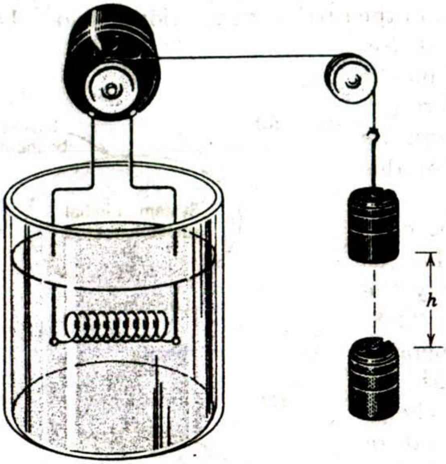


**Fig. 22–8** Heat and work. A weight, in falling, does work on an electric generator which sends current through a resistor which heats the water in which it is immersed.

with the piston. Consider a process whereby the system interacts with its environment and reaches a final equilibrium state characterized by a pressure $p_f$ and a volume $V_f$.

In Fig. 22–9 we show the gas expanding against the piston. The work done by the gas in displacing the piston through an infinitesimal distance $s$ is

$$dW = \mathbf{F} \cdot d\mathbf{s} = pA\,ds = p\,dV$$

where $dV$ is the differential change in the volume of the gas. In general, the pressure will not be constant during a displacement. To obtain the total work $W$ done on the piston by the gas in a large displacement, we must know how $p$ varies with the displacement. Then we compute the integral

$$W = \int dW = \int_{V_i}^{V_f} p\,dV$$

over the range in volume. This integral can be evaluated graphically as the area under the curve in a $p$–$V$ diagram, as shown for a special case in Fig. 22–10.

There are many different ways in which the system can be taken from

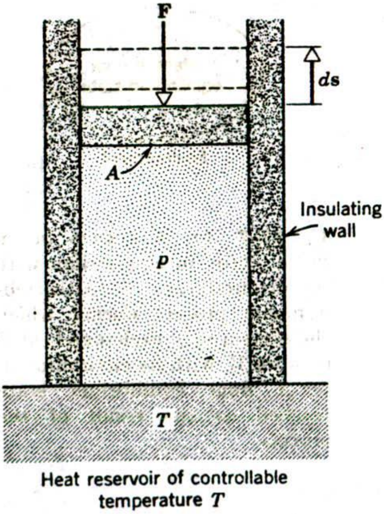


**Fig. 22–9** Work is done by the gas at pressure $p$ as it expands against the piston. Heat may enter or leave the system from the heat reservoir on which the cylinder rests.

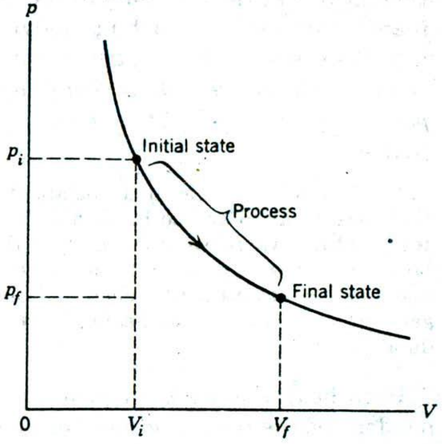


**Fig. 22-10** The work done by a gas is equal to the area under a *p-V* curve.

the initial state *i* to the final state *f*. For example (Fig. 22-11), the pressure may be kept constant from *i* to *a* and then the volume kept constant from *a* to *f*. Then the work done by the expanding gas is equal to the area under the line *ia*. Another possibility is the path *ibf*, in which case the work done by the gas is the area under the line *bf*. The continuous curve from *i* to *f* is another possible path in which the work done by the gas is still different from the previous two paths. We can see, therefore, that *the work done by a system depends not only on the initial and final states but also on the intermediate states, that is, on the path of the process.*

A similar result follows if we compute the flow of heat during the process. State *i* is characterized by a temperature $T_i$ and state *f* by a temperature $T_f$. The heat flowing into the system, say, depends on how the system is heated. We can heat it at a constant pressure $p_i$, for example, until we reach the temperature $T_f$, and then change the pressure at constant temperature to the final value $p_f$. Or we can first lower the pressure to $p_f$ and

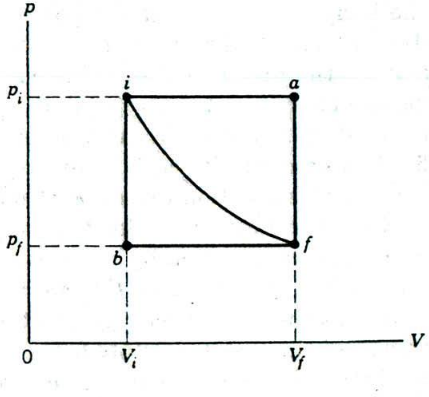


**Fig. 22-11** The work done by a system depends not only on the initial state (*i*) and the final state (*f*) but on the intermediate path as well.

then heat it at that pressure to the final temperature $T_f$. Or we can follow many other paths. Each path gives a different result for the heat flowing into the system. Hence, *the heat lost or gained by a system depends not only on the initial and final states but also on the intermediate states, that is, on the path of the process.* This is an experimental fact. As J. C. Slater has written:

> ". . . It would be pleasant to be able to say, in a given state of the system, that the system has so and so much heat energy. Starting from the absolute zero of temperature, where we could say that the heat energy was zero, we could heat the body up to the state we were interested in, find $\int dQ$ from absolute zero up to this state, and call that the heat energy. But the stubborn fact remains that we would get different answers if we heated it up in different ways. . . . There is nothing to do about it."

Both heat and work "depend on the path" taken; neither one is independent of the path, and neither one can be conserved alone.

## 22–7 The First Law of Thermodynamics

We can now tie all these ideas together. Let a system change from an initial equilibrium state $i$ to a final equilibrium state $f$ in a definite way, the heat absorbed by the system being $Q$ and the work done by the system being $W$. Then we compute the $Q - W$. Now we start over and change the system from the same state $i$ to the same state $f$, but this time in another way by a different path. We do this over and over again, using different paths each time. We find that in every case the quantity *$Q - W$ is the same.* That is, although $Q$ and $W$ separately depend on the path taken, $Q - W$ does not depend at all on how we took the system from state $i$ to state $f$ but only on the initial and final (equilibrium) states.

The student will recall from mechanics that when an object is moved from an initial point $i$ to a final point $f$ in a gravitational field in the absence of friction, the work done depends only on the positions of the two points and not at all on the path through which the body is moved. From this we concluded that there is a function of the space coordinates of the body whose final value minus its initial value equals the work done in displacing the body. We called it the potential energy function. Now in thermodynamics we find that when a system has its state changed from state $i$ to state $f$, the quantity $Q - W$ depends only on the initial and final coordinates and not at all on the path taken between these end points. We conclude that there is a function of the thermodynamic coordinates whose final value minus its initial value equals the change $Q - W$ in the process. We call this function the *internal energy function.*

Now $Q$ is the energy added to the system by the transfer of heat and $W$ is the energy given up by the system in performing work, so that $Q - W$ represents, by definition, *the internal energy change of the system.* Let us represent the internal energy function by the letter $U$. Then the internal energy of the system in state $f$, $U_f$, minus the internal energy of the system in state $i$, $U_i$, is simply *the change in internal energy of the system,* and this quantity *has a definite value independent of how the system went from state i to state f.* We have

$$U_f - U_i = \Delta U = Q - W. \tag{22–6}$$

Just as for potential energy, so for internal energy too it is the change that matters. If some arbitrary value is chosen for the internal energy in some standard reference state, its value in any other state can be given a definite value. Equation 22–6 is known as the *first law of thermodynamics.* In applying Eq. 22–6 we must remember that $Q$ is considered positive when heat enters the system and $W$ is positive when work is done by the system.

If our system undergoes only an infinitesimal change in state, only an infinitesimal amount of heat $dQ$ is absorbed and only an infinitesimal amount of work $dW$ is done, so that the internal energy change $dU$ is also infinitesimal. In such a case, the first law is written in *differential* form as*

$$dU = dQ - dW. \tag{22–7}$$

We may express the first law in words by saying *Every thermodynamic system in an equilibrium state possesses a state variable called the internal energy U whose change dU in a differential process is given by Eq. 22–7.* Recall that the essential content of the zeroth law of thermodynamics (p. 526) is, speaking loosely: *there exists a useful thermodynamic quantity called "temperature."* The essential content of the first law is: *there exists a useful thermodynamic quantity called "internal energy";* the law also provides, in Eq. 22–6, a recipe for measuring changes in internal energy quantitatively.

The first law of thermodynamics is thought to apply to every process in nature that proceeds between equilibrium states. Note that the *process* may or may not involve equilibrium states. We may apply the first law to the explosion of a firecracker in an insulated steel drum, for example. Because of its generality, the information that the first law gives is far from complete, although exact and correct. There are some very general questions which it cannot answer. For example, although it tells us that energy is conserved in every process, it does not tell us whether any particular process can actually occur. An entirely different generalization, called the second law of thermodynamics, gives us this information, and much of the subject matter of thermodynamics depends on this second law (Chapter 25).

## 22–8 Some Applications of the First Law of Thermodynamics

We have seen that when a gas expands the work it does on its environment is

$$W = \int p\,dV,$$

---

* $W$ and $Q$ are not actual functions of the state of a system, that is, they do not depend on the values of the system's coordinates. Hence, $dW$ and $dQ$ are not exact differentials as the term is used in mathematics. All they mean here is a very small quantity. More advanced books write them as $\text{đ}Q$ and $\text{đ}W$ to indicate their inexact nature. However, $dU$ *is* an exact differential, for $U$ is an exact function of the system's coordinates.

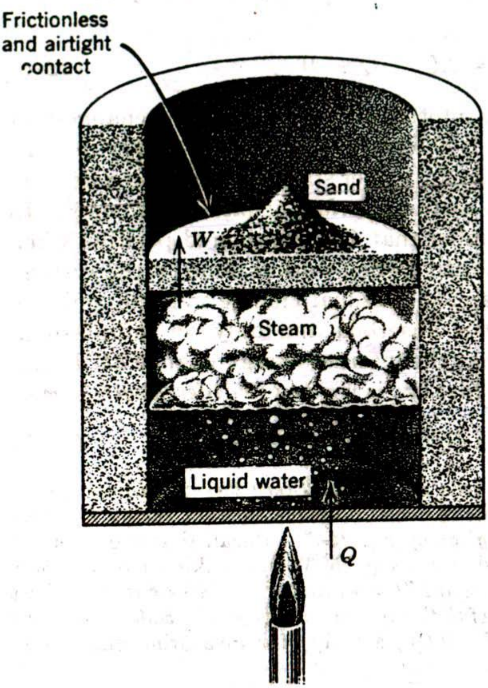


**Fig. 22–12** Water boiling at constant pressure (isobarically). The pressure is kept constant by the weight of the sand and the piston.

where $p$ is the pressure exerted on or by the gas and $dV$ is the differential change in volume of the gas. Consider a special case in which the pressure remains constant while the volume changes by a finite amount, say from $V_i$ to $V_f$. Then

$$W = \int_{V_i}^{V_f} p\,dV = p\int_{V_i}^{V_f} dV = p(V_f - V_i) \qquad \text{(constant pressure).}$$

A process taking place at constant pressure is called an *isobaric* process. For example, water is heated in the boiler of a steam engine up to its boiling point and is vaporized to steam; then the steam is superheated, all processes proceeding at a constant pressure.

In Fig. 22–12 we show an isobaric process. The system is $H_2O$ in a cylindrical container. A frictionless airtight piston is loaded with sand to produce the desired pressure on the $H_2O$ and to maintain it automatically. Heat can be transferred from the environment to the system by a Bunsen burner. If the process continues long enough, the water boils and some is converted to steam; we assume that this occurs. The system may expand, very slowly (quasi-statically) but the pressure it exerts on the piston is automatically always the same, for this pressure must be equal to the constant pressure which the piston exerts on the system. If we wedged the piston so that it could not move, or if we added or took away some sand during the heating process, the process would not be isobaric.

Let us consider the boiling process. We know that substances will change their phase from liquid to vapor at a definite combination of values of pressure and temperature. Water will vaporize at 100° C and atmospheric pressure, for example. For a system to undergo a change of phase heat must be added to it, or taken from it, quite apart from the heat necessary to bring its temperature to the required value. Consider the change of phase of a mass $m$ of liquid to a vapor occurring at constant temperature and pressure. Let $V_l$ be the volume of liquid and $V_v$ the volume of vapor. The work done by this substance in expanding from $V_l$ to $V_v$ at constant pressure is

$$W = p(V_v - V_l).$$

Let $L$ represent the heat of vaporization, that is, the heat needed per unit mass to change a substance from liquid to vapor at constant temperature and pressure. Then the heat absorbed by the mass $m$ during the change of state is

$$Q = mL.$$

From the first law of thermodynamics, we have

$$\Delta U = Q - W$$

so that

$$\Delta U = mL - p(V_v - V_l)$$

for this process.

▶ **Example 3.** At atmospheric pressure 1.00 gm of water, having a volume of 1.00 $\text{cm}^3$, becomes 1671 $\text{cm}^3$ of steam when boiled. The heat of vaporization of water is 539 cal/gm at 1 atm. Hence, if $m = 1.00$ gm,

$$Q = mL = 539 \text{ cal}.$$

This quantity, which respresents heat *added* to the system from the environment, is positive.

$$W = p(V_v - V_l) = (1.013 \times 10^5 \text{ nt/meter}^2)[(1671 - 1) \times 10^{-6} \text{ meter}^3]$$
$$= 169.5 \text{ joules}.$$

This quantity, which represents work done *by* the system on the environment, is positive.

Since 1 cal equals 4.186 joules, $W = 41$ cal. Then,

$$\Delta U = U_v - U_l = mL - p(V_v - V_l) = (539 - 41) \text{ cal}$$
$$= 498 \text{ cal}.$$

This quantity is positive; the internal energy of the system *increases* during this process. Hence, of the 539 cal needed to boil 1 gm of water (at 100° C and 1 atm), 41 cal go into external work of expansion and 498 cal go into internal energy added to the system. This energy represents the internal work done in overcoming the strong attraction of $H_2O$ molecules for one another in the liquid state.

How would you expect the 80 cal that are needed to melt 1 gm of ice to water (at 0° C and 1 atm) to be shared by the external work and the internal energy? ◀

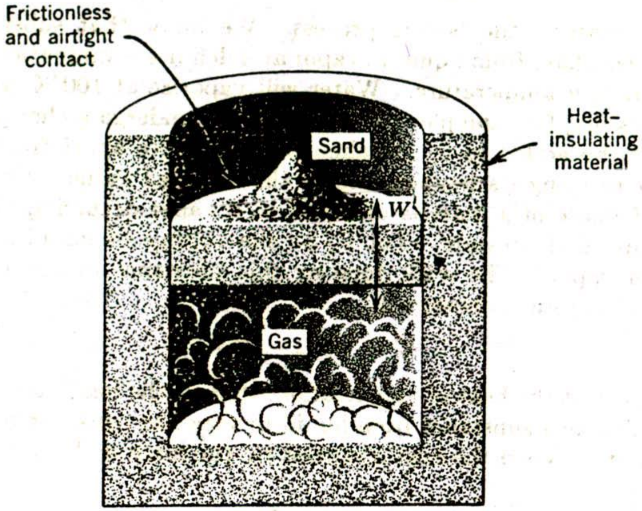


**Fig. 22–13** In an adiabatic process there is no flow of heat to or from the system. Here the walls are insulated and, as sand is removed or added, the volume of the gas changes adiabatically.

A process that takes place in such a way that no heat flows into or out of the system is called an *adiabatic process.* Experimentally such processes are achieved either by sealing the system off from its surroundings with heat insulating material or by performing the process quickly. Because the flow of heat is somewhat slow, any process can be made practically adiabatic if it is performed quickly enough.

For an adiabatic process $Q$ equals zero, so that from the first law we obtain

$$\Delta U = U_f - U_i = -W.$$

Hence, the internal energy of a system increases exactly by the amount of work done *on* the system in an adiabatic process. If work is done *by* the system in an adiabatic process, the internal energy of the system decreases by exactly the amount of external work it performs. An increase of internal energy usually raises the system's temperature and conversely, a decrease of internal energy usually lowers the system's temperature. A gas that expands adiabatically does external work and its internal energy decreases; such a process is used to attain low temperatures. The increase of temperature during an adiabatic compression of air is well known from the heating of a bicycle pump.

In Fig. 22–13 we show a simple adiabatic process. The system is a gas inside a cylinder made of heat-insulating material. Heat cannot enter the system from its environment or leave the system to the environment.

Again we have a pile of sand on a frictionless airtight piston. The only interaction permitted between system and environment is through the performance of work. Such a process can occur when sand is added or removed from the piston, so that the gas can be compressed or can expand against the piston.

Among the many engineering examples of adiabatic processes are the expansion of steam in the cylinder of a steam engine, the expansion of hot gases in an internal combustion engine, and the compression of air in a Diesel engine or in an air compressor. These processes all take place rapidly enough so that only a very small amount of heat can enter or leave the system through its walls during that short time. The compressions and rarefactions in a sound wave are so rapid that the behavior of the transmitting gas is adiabatic (Example 6, Chapter 23).

The most important reason for studying adiabatic processes, however, is that ideal engines use processes that are exactly adiabatic. These ideal engines determine the theoretical limits to the operation and capabilities of real engines. We shall look further into this in Chapter 25.

A process of much theoretical interest is that of *free expansion.* This is an adiabatic process in which no work is performed on or by the system. Something like this can be achieved by connecting one vessel which contains a gas to another evacuated vessel with a stopcock connection, the whole system being enclosed with thermal insulation (Fig. 22-14). If the stopcock is suddenly opened, the gas rushes into the vacuum and expands freely. Because of the heat insulation this process is adiabatic, and because the walls of the vessels are rigid no external work is done on the system. Hence, in the first law we have $Q = 0$ and $W = 0$, so that $U_i = U_f$ for this process. The initial and final internal energies are equal in free expansion.

A free expansion differs from the other examples that we have given in that there is no way to carry it out very slowly (quasi-statically). After

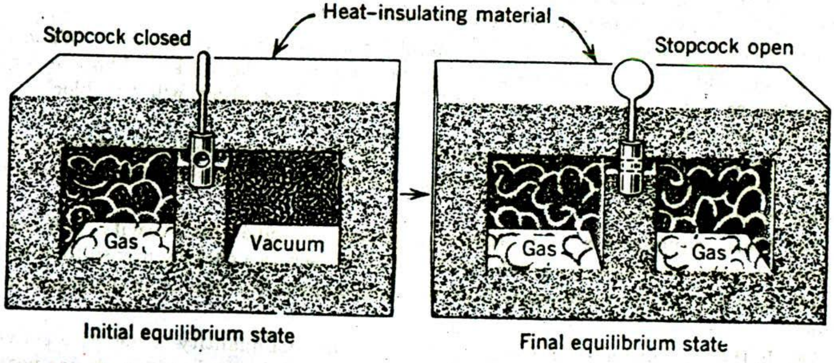


**Fig. 22-14** Free expansion. There is no change of internal energy $U$ since there is no flow of heat $Q$ and no external work $W$ is done.

we open the stopcock we have no further control over the process. At intermediate states the pressure, volume and temperature do not have unique values characteristic of the system as a whole, that is, the system passes through non-equilibrium states so that we cannot plot the course of the process by a curve on a $p$-$V$ diagram. We can plot the initial and final states as points on such plots because they are well-defined, equilibrium states. The free expansion is a good example of an irreversible process; see Section 25–2.

## QUESTIONS

**1.** Give examples to distinguish clearly between temperature and heat.

**2.** (*a*) Show how heat conduction and calorimetry could be explained by the caloric theory. (*b*) List some heat phenomena that cannot be explained by the caloric theory.

**3.** Can heat be considered a form of stored (or potential) energy? Would such an interpretation contradict the concept of heat as energy in process of transfer because of a temperature difference?

**4.** Apply Eq. 22–1 to boiling water.

**5.** Can heat be added to a substance without causing the temperature of the substance to rise? If so, does this contradict the concept of heat as energy in process of transfer because of a temperature difference?

**6.** Explain the fact that the presence of a large body of water nearby, such as a sea or ocean, tends to moderate the temperature extremes of the climate on adjacent land.

**7.** Theory shows that the coefficient of linear expansion $\alpha$ (see Sec. 21–8) is proportional to the heat capacity $C_v$. Show that this is to be expected. (Hint: heat capacity measures the rate of change of the vibrational energy with temperature.)

**8.** Give an example of a process in which no heat is transferred to or from the system but the temperature of the system changes.

**9.** Both heat conduction and wave propagation involve the transfer of energy. Is there any difference in principle between these two phenomena?

**10.** When a hot body warms a cool one are their temperature changes equal in magnitude? Give examples. Can one then say that temperature passes from one to the other?

**11.** What connection is there between an object's feeling hot or cold and its heat capacity? Between this and its thermal conductivity?

**12.** A block of wood and a block of metal are at the *same* temperature. When the blocks feel cold the metal feels colder than the wood; when the blocks feel hot the metal feels hotter than the wood. Explain. At what temperature will the blocks feel equally cold or hot?

**13.** On a winter day the temperature of the inside surface of a wall is much lower than room temperature and that of the outside surface is much higher than the outdoor temperature. Explain.

**14.** What requirements for thermal conductivity, specific heat capacity, and coefficient of expansion would you want a material to be used as a cooking utensil to satisfy?

**15.** In what way is steady-state heat flow analogous to the flow of an incompressible fluid?

**16.** Is the mechanical equivalent of heat, $J$, a physical quantity or merely a conversion factor for converting energy from heat units to mechanical units and vice versa?

**17.** Is the temperature of an isolated system (no interaction with the environment) conserved?

**18.** Can one distinguish between whether the internal energy of a body was acquired by heat transfer or acquired by performance of work?

**19.** If the pressure and volume of a system are given is the temperature always uniquely determined?

**20.** Does a gas do any work when it expands adiabatically? If so, what is the source of the energy needed to do this work?

**21.** A quantity of gas occupies an initial volume $V_0$ at a pressure $p_0$ and a temperature $T_0$. It expands to a volume $V$ (a) at constant temperature and (b) at constant pressure. In which case does the gas do more work?

**22.** Discuss the process of the freezing of water from the point of view of the first law of thermodynamics. Remember that ice occupies a greater volume than an equal mass of water.

**23.** A thermos bottle contains coffee. The thermos bottle is vigorously shaken. Consider the coffee as the system. (*a*) Does its temperature rise? (*b*) Has heat been added to it? (*c*) Has work been done on it? (*d*) Has its internal energy changed?

## PROBLEMS

**1.** In a Joule experiment, a mass of 6.00 kg falls through a height of 50.0 meters and rotates a paddle wheel which stirs 0.600 kg of water. The water is initially at 15° C. By how much does its temperature rise?

**2.** Compute the possible increase in temperature for water going over Niagara Falls, 162 ft high. What factors would tend to prevent this possible rise?

**3.** An energetic athlete dissipates all the energy in a diet of 4000 kcal per day. If he were to release this heat at a steady rate, how would his heat output compare with the energy output of a 100-watt bulb?

**4.** A block of ice at 0° C whose mass is initially 50.0 kg slides along a horizontal surface, starting at a speed of 5.38 meters/sec and finally coming to rest after traveling 28.3 meters. Compute the mass of ice melted as a result of the friction between the block and the surface. List any assumptions you need to make in getting your answer.

**5.** Calculate the specific heat of a metal from the following data. A container made of the metal weighs 8.0 lb and contains in addition 30 lb of water. A 4.0-lb piece of the metal initially at a temperature of 350° F is dropped into the water. The water and container initially have a temperature of 60° F and the final temperature of the entire system is 65° F.

**6.** A thermometer of mass 0.055 kg and of specific heat 0.20 kcal/kgC° reads 15.0° C. It is then inserted into 0.300 kg of water and it comes to the same final temperature of the water. If the thermometer reads 44.4° C and is accurate, what was the temperature of the water before insertion of the thermometer, neglecting other heat losses?

**7.** Count Rumford weighed a metal object at low temperature and then at high temperature to see whether its "caloric content" increased. He concluded that (for gold) the "caloric" did not weigh more than $10^{-6}$ the weight of the sample. (*a*) Should the mass of a sample increase when heated, according to modern theories? (*b*) If so, by what order of magnitude? (c) Was Rumford safe in rejecting the caloric theory on this basis, in retrospect?

**8.** Take the average specific heat of copper to be 0.090 cal/gmC° in the temperature range 0 to 1000° C. If 1.00 kg of copper is heated from 0 to 1000° C. by how much does its mass increase?

**9.** A "flow calorimeter" is used to measure the specific heat of a liquid. Heat is added at a known rate to a stream of the liquid as it passes through the calorimeter at a known rate. Then a measurement of the resulting temperature difference between the inflow and the outflow points of the liquid stream enables us to compute the specific heat of the liquid.

A liquid of density 0.85 gm/cm$^3$ flows through a calorimeter at the rate of 8.0 cm$^3$/sec. Heat is added by means of a 250-watt electric heating coil, and a temperature difference of 15 C° is established in steady-state conditions between the inflow and outflow points. Find the specific heat of the liquid.

**10.** By means of a heating coil energy is transferred at a *constant* rate to a substance in a *thermally insulated* container. The temperature of the substance is measured as a function of the time. Show how we can deduce the way in which the heat capacity of the body depends on the temperature from this information.

**11.** Suppose the specific heat of a substance is found to vary with temperature in a parabolic fashion, that is

$$c = A + BT^2,$$

where $A$ and $B$ are constants and $T$ is Celsius temperature. Compare the *mean* specific heat of the substance in a temperature range $T = 0$ to $T = T$ to the specific heat at the midpoint $T/2$.

**12.** The specific heat of silver, measured at atmospheric pressure, is found to vary with temperature between 50 and 100° K by the empirical equation

$$c_p = 0.076T - 0.00026T^2 - 0.15,$$

where $c_p$ is in cal/mole K° and $T$ is the Kelvin temperature. Calculate the quantity of heat required to raise 216 gm of silver from 50 to 100° K.

**13.** Power is supplied at the rate of 0.40 hp for 2.0 min in drilling a hole in a 1.0-lb brass block. (a) How much heat is generated? (b) What is the rise in temperature of the brass if 75% of the heat generated warms the brass? (c) What happens to the other 25%?

**14.** A 2.0-gm lead bullet moving at a speed of 200 meters/sec becomes embedded in a 2.0-kg wooden block of a ballistic pendulum. Calculate the rise in temperature of the bullet, assuming that all the heat generated raises the bullet's temperature.

**15.** Consider the rod shown in Fig. 22–4. Suppose $L = 25$ cm, $A = 1.0$ cm$^2$, and the material is copper. If $T_2 = 125°$ C, $T_1 = 0°$ C, and a steady state is reached, find (a) the temperature gradient, (b) the rate of heat transfer, and (c) the temperature at a point in the rod 10 cm from the high-temperature end.

**16.** Show that in a compound slab the temperature gradient in each portion is inversely proportional to the thermal conductivity.

**17.** Assume that the thermal conductivity of copper is twice that of aluminum and four times that of brass. Three metal rods, made of copper, aluminum, and brass, respectively, are each 6.0 in. long and 1.0 in. in diameter. These rods are placed end-to-end, with the aluminum between the other two. The free ends of the copper and brass rods are maintained at 100 and 0° C, respectively. Find the equilibrium temperatures of the copper-aluminum junction and the aluminum-brass junction.

**18.** Assuming $k$ is constant, show that the radial rate of flow of heat in a substance between two concentric spheres is given by

$$\frac{dQ}{dt} = \frac{(T_1 - T_2)4\pi k r_1 r_2}{r_2 - r_1}$$

where the inner sphere has a radius $r_1$ and temperature $T_1$, and the outer sphere has a radius $r_2$ and temperature $T_2$.

**19.** Heat generated by radioactivity within the earth is conducted outward through the oceans. For purposes of approximate calculation, assume the average temperature gradient within the solid earth beneath the ocean to be 0.07 C°/meter and the average thermal conductivity to be $2 \times 10^{-4}$ kcal/meter sec C°, and determine the rate of heat transfer per square meter. Assume that this is approximately the rate for the entire surface of the earth, and determine how much heat is thereby transferred through the earth's surface each day.

**20.** Assuming $k$ is constant, show that the radial rate of flow of heat in a substance between two coaxial cylinders is given by

$$\frac{dQ}{dt} = \frac{(T_1 - T_2)2\pi Lk}{\ln (r_2/r_1)}$$

where the inner cylinder has a radius $r_1$ and temperature $T_1$, and the outer cylinder has a radius $r_2$ and temperature $T_2$, each cylinder having a length $L$.

**21.** A long tungsten heater wire is rated at 3.0 kw/meter and is $5.0 \times 10^{-4}$ meter in diameter. It is embedded along the axis of a ceramic cylinder of diameter 0.12 meter. When operating at the rated power, the wire is at 1500° C; the outside of the cylinder is at 20° C. Find the thermal conductivity of the ceramic; use the result given in Problem 20.

**22.** Determine the value of $J$, the mechanical equivalent of heat, from the following data: 2000 cal of heat are supplied to a system; the system does 3350 joules of external work during that time; the increase in internal energy during the process is 5030 joules.

**23** When a system is taken from state $i$ to state $f$ along the path $iaf$, it is found that 50 cal and $W = 20$ cal. Along the path $ibf$, $Q = 36$ cal (Fig. 22–15). (*a*) What is $W$ along the path $ibf$? (*b*) If $W = -13$ cal for the curved return path $fi$, what is $Q$ for this path? (*c*) Take $U_i = 10$ cal. What is $U_f$? (*d*) If $U_b = 22$ cal, what is $Q$ for the process $ib$? For the process $bf$?

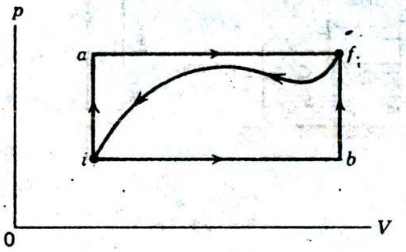
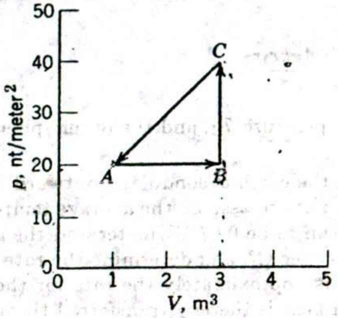


Fig. 22–15

**24.** A thermodynamic system is taken from an initial state $A$ to another $B$ and back again to $A$, via state $C$, as shown by the path $A$-$B$-$C$-$A$ in the $p$-$V$ diagram of Fig. 22–16*a*. (*a*) Complete the table in Fig. 22–16*b* by filling in appropriate + or − indi-



(a)

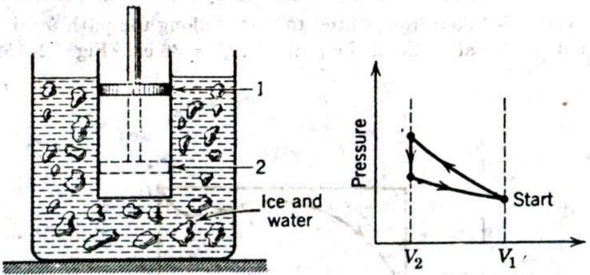
|  | $Q$ | $W$ | $\Delta U$ |
|---|---|---|---|
| $A \longrightarrow B$ |  |  | + |
| $B \longrightarrow C$ | + |  |  |
| $C \longrightarrow A$ |  |  |  |

(b)

Fig. 22–16

cations for the signs of the thermodynamic quantities associated with each process. (*b*) Calculate the numerical value of the work done by the system for the complete cycle *A-B-C-A*.

**25.** Figure 22–17*a* shows a cylinder containing gas and closed by a movable piston. The cylinder is submerged in an ice-water mixture. The piston is *quickly* pushed down from position (1) to position (2). The piston is held at position (2) until the gas is again at 0° C and then is *slowly* raised back to position (1). Figure 22–17*b* is a *p*-*V* diagram for the process. If 100 gm of ice are melted during the cycle, how much work has been done *on* the gas?



(a) (b)

Fig. 22–17

**26.** An iron ball is dropped onto a concrete floor from a height of 10 meters. On the first rebound it rises to a height of 0.50 meter. Assume that all the macroscopic mechanical energy lost in the collision with the floor goes into the ball. The specific heat of iron is 0.12 cal/gm C°. During the collision (*a*) has heat been added to the ball? (*b*) Has work been done on it? (*c*) Has its internal energy changed? If so, by how much? (*d*) How much has the temperature of the ball risen after the first collision?

# Kinetic Theory of Gases—I

## CHAPTER 23

### 23–1 Introduction

*Thermodynamics* deals only with macroscopic variables, such as pressure, temperature, and volume. Its basic laws, expressed in terms of such quantities, say nothing at all about the fact that matter is made up of atoms. *Statistical mechanics,* however, which deals with the same areas of science that thermodynamics does, presupposes the existence of atoms. Its basic laws are the laws of mechanics, which are applied to the atoms that make up the system.

No existing electronic computer could solve the problem of applying the laws of mechanics individually to every atom in a gas, say. If there were one, the results of such calculations would be too voluminous to be useful. Fortunately, the detailed life histories of individual atoms in a gas are not important if we want to calculate only the macroscopic behavior of the gas. We apply the laws of mechanics *statistically*, then, and we find that we are able to express all the thermodynamic variables as certain averages of atomic properties. For example, the pressure exerted by a gas on the wall of the containing vessel is the average rate per unit area at which the atoms of the gas transfer momentum to the wall as they collide with it. The number of atoms in a macroscopic system is usually so large that such averages are very sharply defined quantities indeed.

We can apply the laws of mechanics statistically to assemblies of atoms at two different levels. At the level called *kinetic theory* we proceed in a rather physical way, using relatively simple mathematical averaging techniques. In this chapter we will use these methods to enlarge our understanding of pressure, temperature, specific heat, and internal energy at the atomic level. Kinetic theory was developed by Robert Boyle (1627–1691), Daniel Bernoulli (1700–1782), James Joule (1818–1889), A. Kronig (1822–1879), Rudolph Clausius (1822–1888), and Clerk Maxwell (1831–1879), among others.* In this book we apply the kinetic theory to gases only, because the interactions between atoms in gases are much weaker than in liquids and solids; this greatly simplifies the mathematical difficulties.

At another level, we can apply the laws of mechanics statistically using techniques that are more formal and abstract than those of kinetic theory. This approach, developed by J. Willard Gibbs (1839–1903) and by Ludwig Boltzmann (1844–1906) among others, is called *statistical mechanics*, a term that includes kinetic theory as a sub-branch. Using these methods one can derive the laws of thermodynamics, thus establishing that science as a branch of mechanics. The fullest flowering of statistical mechanics (*quantum statistics*) involves the statistical application of the laws of quantum mechanics—rather than those of classical mechanics—to many-atom systems.†

## 23–2 Ideal Gas—A Macroscopic Description

Let a mass $\mathfrak{M}$ of a gas be confined in a container of volume $V$. The density $\rho$ of the gas is $\mathfrak{M}/V$ and it is clear that we can reduce $\rho$ either by removing some gas from the container (reducing $\mathfrak{M}$) or by putting the gas in a larger container (increasing $V$). We find from experiment that, at low enough densities, all gases, no matter what their chemical composition, tend to show a certain simple relationship among the thermodynamic variables $p$, $V$, and $T$. This suggests the concept of an *ideal gas*, one that would have the same simple behavior under all conditions. In this section we give a macroscopic or thermodynamic definition of an ideal gas. In Section 23–3 we will define an ideal gas microscopically, from the standpoint of kinetic theory, and we will see what we can learn by comparing these two approaches.

Given a mass $\mathfrak{M}$ of any gas in a state of thermal equilibrium we can measure its pressure $p$, its temperature $T$, and its volume $V$. For low enough values of the density experiment shows that (1) for a given mass of gas held at a constant temperature, the pressure is inversely proportional to the volume (Boyle's law), and (2) for a given mass of gas held at a constant pressure, the volume is directly proportional to the temperature (law of Charles and Gay-Lussac). We can summarize these two experimental results by the relation

$$\frac{pV}{T} = \text{a constant} \quad \text{(for a fixed mass of gas).} \tag{23–1}$$

---

* See "John James Waterston and the Kinetic Theory of Gases," by S. G. Brush, in *American Scientist*, June 1961, for an interesting aspect of the history of kinetic theory.

† See *Thermal Physics* by Philip M. Morse, W. A. Benjamin, Inc., New York, 1962, for a fuller treatment of thermodynamics, kinetic theory, and (particularly) statistical mechanics proper, than we can give here.

The volume occupied by a gas (real or ideal) at a given pressure and temperature is proportional to its mass. Thus the constant in Eq. 23–1 must also be proportional to the mass of the gas. In Section 22–2 (see Fig. 22–2) we saw the great simplification that occurs in studies of the specific heats of solids if we compare samples of solids that contain the same number of molecules rather than samples that have the same mass in grams. We did this by using the mole as our mass unit. Let us also do that here.

We therefore write the constant in Eq. 23–1 as $\mu R$, where $\mu$ is the mass of the gas in moles and $R$ is a constant that must be determined for each gas by experiment. Our expectation that simplicity will emerge if we compare gases on a molar basis is justified because experiment shows that, at low enough densities, *$R$ has the same value for all gases*, namely

$$R = 8.314 \text{ joule/mole K}^\circ = 1.986 \text{ cal/mole K}^\circ.$$

$R$ is called the *universal gas constant*. We then write Eq. 23–1 as

$$pV = \mu RT \tag{23-2}$$

and we define an ideal gas as one that obeys this relation *under all conditions*. There is no such thing as a truly ideal gas, but it remains a useful and simple concept connected with reality by the fact that all real gases approach the ideal gas abstraction in their behavior if the density is low enough. Equation 23–2 is called the *equation of state* of an ideal gas.

If we could fill the bulb of an (ideal) constant-volume gas thermometer with an ideal gas, we see from Eq. 23–2 that we could define temperature in terms of its pressure readings, that is,

$$T = 273.16^\circ \text{ K} \frac{p}{p_{tr}} \quad \text{(ideal gas)}.$$

Here $p_{tr}$ is the gas pressure at the triple point, at which the temperature $T_{tr}$ is 273.16° K by definition. In practice we must fill our thermometer with a real gas and measure the temperature by extrapolating to zero density using Eq. 21–4,

$$T = 273.16^\circ \text{ K} \lim_{p_{tr}\to 0} \frac{p}{p_{tr}} \quad \text{(real gas)}.$$

▶ **Example 1.** A cylinder contains oxygen at a temperature of 20° C and a pressure of 15 atm in a volume of 100 liters. A piston is lowered into the cylinder decreasing the volume occupied by the gas to 80 liters and raising the temperature to 25° C. Assuming oxygen to behave like an ideal gas under these conditions, what then is the gas pressure?

From Eq. 23–1, since the mass of gas remains unchanged, we may write

$$\frac{p_iV_i}{T_i} = \frac{p_fV_f}{T_f}.$$

Our initial conditions are

$$p_i = 15 \text{ atm}, \qquad T_i = 293^\circ \text{ K}, \qquad V_i = 100 \text{ liters}.$$

Our final conditions are

$$p_f = ?, \qquad T_f = 298^\circ \text{ K}, \qquad V_f = 80 \text{ liters}.$$

Hence,

$$p_f = \left(\frac{T_f}{V_f}\right)\left(\frac{p_i V_i}{T_i}\right) = \left(\frac{298^\circ \text{ K}}{80 \text{ liters}}\right)\left(\frac{15 \text{ atm} \times 100 \text{ liters}}{293^\circ \text{ K}}\right) = 19 \text{ atm}.$$

▶ **Example 2.** Calculate the work per mole done by an ideal gas which expands isothermally, that is, at constant temperature, from an initial volume $V_i$ to a final volume $V_f$.

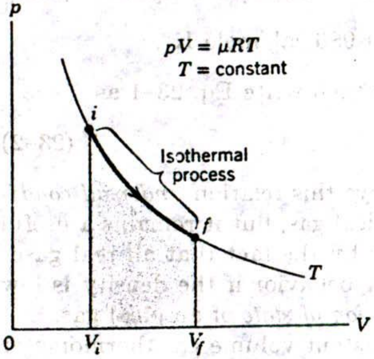


**Fig. 23–1** Example 2. The shaded area represents the work done by $\mu$ moles of gas in expanding from $V_i$ to $V_f$ with the temperature held constant.

The work done may be represented as

$$W = \int_{V_i}^{V_f} p\, dV.$$

From the ideal gas law we have

$$p = \frac{\mu RT}{V},$$

so that $W/\mu$, the work per mole, is

$$\frac{W}{\mu} = \int_{V_i}^{V_f} \frac{RT}{V}\, dV.$$

The temperature is constant so that

$$\frac{W}{\mu} = RT \int_{V_i}^{V_f} \frac{dV}{V} = RT \ln \frac{V_f}{V_i}$$

is the work per mole done by an ideal gas in an isothermal expansion at temperature $T$ from an initial volume $V_i$ to a final volume $V_f$.

Notice that when the gas expands, so that $V_f > V_i$, the work done by the gas is positive; when the gas is compressed, so that $V_f < V_i$, the work done by the gas is negative. This is consistent with the sign convention adopted for $W$ in the first law of thermodynamics. The work done is shown as the shaded area in Fig. 23–1. The solid line is an isotherm, that is, a curve giving the relation of $p$ to $V$ at a constant temperature.

In practice, how can we keep an expanding or contracting gas at constant temperature? ◀

## 23–3 An Ideal Gas—Microscopic Definition

From the microscopic point of view we define an ideal gas by making the following assumptions; it will then be our task to apply the laws of classical mechanics statistically to the gas atoms and to show that our microscopic definition is consistent with the macroscopic definition of the preceding section:

1. *A gas consists of particles called molecules.* Depending on the gas, each molecule will consist of one atom or a group of atoms. If the gas is an element or a compound and is in a stable state, we consider all its molecules to be identical.

2. *The molecules are in random motion and obey Newton's laws of motion.* The molecules move in all directions and with various speeds. In computing the properties of the motion, we assume that Newtonian mechanics works at the microscopic level. As for all our assumptions, this one will stand or fall depending on whether or not the experimental facts it predicts are correct.

3. *The total number of molecules is large.* The direction and speed of motion of any one molecule may change abruptly on collision with the wall or another molecule. Any particular molecule will follow a zigzag path because of these collisions. However, because there are so many molecules we assume that the resulting large number of collisions maintains the over-all distribution of molecular velocities and the randomness of the motion.

4. *The volume of the molecules is a negligibly small fraction of the volume occupied by the gas.* Even though there are many molecules, they are extremely small. We know that the volume occupied by a gas can be changed through a large range of values with little difficulty, and that when a gas condenses the volume occupied by the liquid may be thousands of times smaller than that of the gas. Hence, our assumption is plausible. Later we shall investigate the actual size of molecules and see whether we need to modify this assumption.

5. *No appreciable forces act on the molecules except during a collision.* To the extent that this is true a molecule moves with uniform velocity between collisions. Because we have assumed the molecules to be so small, the average distance between molecules is large compared to the size of a molecule. Hence, we assume that the range of molecular forces is comparable to the molecular size.

6. *Collisions are elastic and are of negligible duration.*. Collisions between molecules and with the walls of the container conserve momentum and (we assume) kinetic energy. Because the collision time is negligible compared to the time spent by a molecule between collisions, the kinetic energy which is converted to potential energy during the collision is available again as kinetic energy after such a brief time that we can ignore this exchange entirely.

## 23-4 Kinetic Calculation of the Pressure

Let us now calculate the pressure of an ideal gas from kinetic theory. To simplify matters, we consider a gas in a cubical vessel whose walls are perfectly elastic. Let each edge be of length $l$. Call the faces normal to the $x$-axis (Fig. 23-2) $A_1$ and $A_2$, each of area $l^2$. Consider a molecule which has a velocity $\mathbf{v}$. We can resolve $\mathbf{v}$ into components $v_x$, $v_y$, and $v_z$ in the directions of the edges. If this particle collides with $A_1$, it will

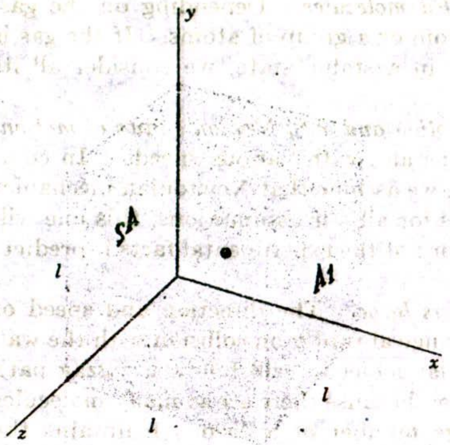


**Fig. 23–2** A cubical box of side $l$, containing an ideal gas. A molecule is shown moving toward $A_1$.

rebound with its $x$-component of velocity reversed. There will be no effect on $v_y$ or $v_z$, so that the change $\Delta p$ in the particle's momentum will be

$$\Delta p = p_f - p_i = -mv_x - (mv_x) = -2mv_x,$$

normal to $A_1$. Hence, the momentum imparted to $A_1$ will be $2mv_x$, since the total momentum is conserved.

Suppose that this particle reaches $A_2$ without striking any other particle on the way. The time required to cross the cube will be $l/v_x$. At $A_2$ it will again have its $x$-component of velocity reversed and will return to $A_1$. Assuming no collisions in between, the round trip will take a time $2l/v_x$. Hence, the number of collisions per unit time this particle makes with $A_1$ is $v_x/2l$, so that the rate at which it transfers momentum to $A_1$ is

$$2mv_x \frac{v_x}{2l} = \frac{mv_x^2}{l}.$$

To obtain the total force on $A_1$, that is, the rate at which momentum is imparted to $A_1$ by all the gas molecules, we must sum up $mv_x^2/l$ for all the particles. Then, to find the pressure, we divide this force by the area of $A_1$, namely $l^2$.

If $m$ is the mass of each molecule, we have

$$p = \frac{m}{l^3}\,(v_{x1}^2 + v_{x2}^2 + \cdots),$$

where $v_{x1}$ is the $x$-component of the velocity of particle 1, $v_{x2}$ is that of particle 2, etc. If $N$ is the total number of particles in the container and $n$ is the number per unit volume, then $N/l^3 = n$ or $l^3 = N/n$. Hence,

$$p = mn\left(\frac{v_{x1}^2 + v_{x2}^2 + \cdots}{N}\right).$$

But $mn$ is simply the mass per unit volume, that is, the density $\rho$. The quantity $(v_{x1}^2 + v_{x2}^2 + \cdots)/N$ is the average value of $v_x^2$ for all the particles in the container. Let us call this $\overline{v_x^2}$. Then

$$p = \rho\overline{v_x^2}.$$

For any particle $v^2 = v_x^2 + v_y^2 + v_z^2$. Because we have many particles and because they are moving entirely at random, the average values of $v_x^2$, $v_y^2$, and $v_z^2$ are equal and the value of each is exactly one-third the average value of $v^2$. There is no preference among the molecules for motion along any one of the three axes. Hence, $\overline{v_x^2} = \frac{1}{3}\overline{v^2}$, so that

$$p = \rho\overline{v_x^2} = \tfrac{1}{3}\rho\overline{v^2}. \tag{23-3}$$

Although we derived this result by neglecting collisions between particles, the result is true even when we consider collisions. Because of the exchange of velocities in an elastic collision between identical particles, there will always be some one molecule that will collide with $A_2$ with momentum $mv_x$ corresponding to the one that left $A_1$ with this same momentum. Also, the time spent during collisions is negligible compared to the time spent between collisions. Hence, our neglect of collisions is merely a convenient device for calculation. Likewise, we could have chosen a container of any shape—the cube merely simplifies the calculation. Although we have calculated the pressure exerted only on the side $A_1$, it follows from Pascal's law that the pressure is the same on all sides and everywhere in the interior.*

The square root of $\overline{v^2}$ is called the *root-mean-square* speed of the molecules and is a kind of average molecular speed.† Using Eq. 23-3, we can calculate this root-mean-square speed from measured values of the pressure and density of the gas. Thus,

$$v_{\text{rms}} = \sqrt{\overline{v^2}} = \sqrt{\frac{3p}{\rho}}. \tag{23-4}$$

In Eq. 23-3 we relate a macroscopic quantity (the pressure $p$) to an average value of a microscopic quantity (that is, to $\overline{v^2}$ or $v_{\text{rms}}^2$. However, averages can be taken over short times or over long times, over small regions of space or large regions of space. The average computed in a small region for a short time might depend on the time or region chosen, so that the values obtained in this way may fluctuate. This could happen in a gas of very low density, for example. We can ignore fluctuations, however, when the number of particles in the system is large enough.

---

* We neglect the weight of the gas, a negligible effect unless the gas is of very large extent, as in the atmosphere. (See Section 17-3 and Problem 42.)

† We will consider this further in Section 24-2 in which we discuss the molecular distribution of speeds.

▶ **Example 3.** Calculate the root-mean-square speed of hydrogen molecules at 0.00° C and 1.00-atm pressure, assuming hydrogen to be an ideal gas. Under these conditions hydrogen has a density $\rho$ of $8.99 \times 10^{-2}$ kg/meter$^3$. Then, since $p = 1.00 \text{ atm} = 1.01 \times 10^5$ nt/meter$^2$,

$$v_{\text{rms}} = \sqrt{\frac{3p}{\rho}} = 1840 \text{ meters/sec.}$$

This is of the order of a mile per second, or 3600 miles/hr. ◀

Table 23–1 gives the results of similar calculations for some gases at 0° C.

**Table 23–1**

| Gas | $v_{\text{rms}}$ (at 0° C), meters/sec | Molecular weight,* gm/mole | Translational kinetic energy per mole (at 0° C), $\frac{1}{2}Mv_{\text{rms}}^2$ joules/mole |
|---|---|---|---|
| $O_2$ | 461 | 32 | 3400 |
| $N_2$ | 493 | 28 | 3390 |
| Air | 485 | 28.8 | 3280 |
| CO | 493 | 28 | 3390 |
| $H_2$ | 1838 | 2.02 | 3370 |
| He | 1311 | 4.0 | 3430 |
| $CO_2$ | 393 | 44 | 3400 |
| $H_2O$ | 615 | 18 | 3400 |
| Ne | 584 | 20.1 | 3420 |

* The molecular weight and the mole are defined on page 550. We will discuss the last column in this table in the next section.

These molecular speeds are of the same order as the speed of sound at the same temperature. For example, in air at 0° C, $v_{\text{rms}} = 485$ meters/sec and the speed of sound is 331 meters/sec; in hydrogen $v_{\text{rms}} = 1838$ meters/sec and sound travels at 1286 meters/sec; in oxygen $v_{\text{rms}} = 461$ meters/sec and sound travels at 317 meters/sec. These results are to be expected in terms of our model of a gas; see Prob. 34. We visualize the propagation of sound waves as a directional motion of the molecules as a whole superimposed on their random motion. Hence, the energy of the sound wave is carried as kinetic energy from one molecule to the next one with which it collides. The molecules themselves, in spite of their high speeds, do not move very far during a period of the sound vibration; they are confined to a rather small space by the effects of a large number of collisions.*

---

* This explains why there is a time lag between opening an ammonia bottle at one end of the room and smelling it at the other end. Although molecular speeds are high, the large number of collisions restrains the advance of the ammonia molecules. They diffuse through the air at speeds that are rather small compared to molecular speeds.

However, the energy of the sound wave *is* communicated from one molecule to the next with that high speed, even though we do not expect the speed of sound to be *exactly* equal to $v_{\text{rms}}$, a point that we will clarify in Example 6.

▶ **Example 4.** Assume that the speed of sound in a gas is the same as the root-mean-square speed of the molecules, and show how the speed of sound for an ideal gas depends on the temperature.

The density of a gas is

$$\rho = \frac{\mathfrak{M}}{V} = \frac{\mu M}{V}$$

in which $\mathfrak{M}$ is the mass of the gas, $M$ is the molecular weight (grams/mole), and $\mu$ is the mass in moles. Combining this with the ideal gas law

$$pV = \mu RT$$

yields

$$\frac{p}{\rho} = \frac{RT}{M}.$$

We obtain from Eq. 23–4

$$v_{\text{rms}} = \sqrt{\frac{3p}{\rho}} = \sqrt{\frac{3RT}{M}},$$

so that the speed of sound $v_1$ at a temperature $T_1$ is related to the speed of sound $v_2$ in the same gas at a temperature $T_2$ by

$$\frac{v_1}{v_2} = \sqrt{\frac{T_1}{T_2}}.$$

For example, if the speed of sound at 273° K is 332 meters/sec in air, its speed in air at 300° K will be

$$\sqrt{\tfrac{300}{273}} \times 332 \text{ meters/sec} = 348 \text{ meters/sec}.$$

Would our result change if the speed of sound were proportional to, rather than equal to, the root-mean-square speed of the molecules of a gas? ◀

## 23–5 Kinetic Interpretation of Temperature

If we multiply each side of Eq. 23–3 by the volume $V$, we obtain

$$pV = \tfrac{1}{3}\rho V\overline{v^2},$$

where $\rho V$ is simply the total mass $\mathfrak{M}$ of gas, $\rho$ being the density. We can also write the mass of gas as $\mu M$, in which $\mu$ is the mass in moles and $M$ is the molecular weight. Making this substitution yields

$$pV = \tfrac{1}{3}\mu M\overline{v^2}.$$

The quantity $\frac{1}{3}\mu M\overline{v^2}$ is two-thirds the total kinetic energy of translation of the molecules, that is, $\frac{2}{3}(\frac{1}{2}\mu M\overline{v^2})$[*]. We can then write

$$pV = \tfrac{2}{3}(\tfrac{1}{2}\mu M\overline{v^2}).$$

The equation of state of an ideal gas is

$$pV = \mu RT.$$

Combining these two expressions, we obtain

$$\tfrac{1}{2}M\overline{v^2} = \tfrac{3}{2}RT. \tag{23-5}$$

That is, *the total translational kinetic energy per mole of the molecules of an ideal gas is proportional to the temperature.* We may say that this result, Eq. 23–5, is necessary to fit the kinetic theory to the equation of state of an ideal gas, or we may consider Eq. 23–5 as a definition of gas temperature on a kinetic theory or microscopic basis. In either case, we gain some insight into the meaning of temperature for gases.

The temperature of a gas is related to the total translational kinetic energy measured with respect to the center of mass of the gas. The kinetic energy associated with the motion of the center of mass of the gas has no bearing on the gas temperature. In Section 23–3 we assumed random motion as part of our statistical definition of an ideal gas and in Section 23–4 we calculated $\overline{v^2}$ on this basis. For a random distribution of molecular velocities with direction the center of mass would be at rest, so that we must use a reference frame in which the center of mass of the gas is at rest. For all other frames the molecules will each have velocities greater by **u** (the velocity of the center of mass in that frame) than in the center of mass frame; hence, the motions will no longer be random and we will obtain different values for $\overline{v^2}$. The temperature of a gas in a container does not increase when we put the container on a moving train!

Let us now divide each side of Eq. 23–5 by Avogadro's number, $N_0$, which (see page 550, footnote) is the number of molecules per mole of a gas. Thus $M/N_0$ $(= m)$ is the mass of a single molecule and we have

$$\tfrac{1}{2}(M/N_0)\overline{v^2} = \tfrac{1}{2}m\overline{v^2} = \tfrac{3}{2}(R/N_0)T.$$

Now $\frac{1}{2}m\overline{v^2}$ is the average translational kinetic energy per molecule. The ratio $R/N_0$—which we call $k$, the *Boltzmann constant*—plays the role of the gas constant per molecule. We have

$$\tfrac{1}{2}m\overline{v^2} = \tfrac{3}{2}kT \tag{23-6}$$

---

[*] If $N$ is the total number of molecules and $m$ is the mass of each molecule, then $\frac{1}{2}mv_1^2 + \frac{1}{2}mv_2^2 + \cdots = \frac{1}{2}mN\left[\dfrac{v_1^2 + v_2^2 + \cdots}{N}\right] = \frac{1}{2}mN\overline{v^2}$ in which $mN$ $(= \mu M)$ is the total mass of the gas.

in which*

$$k = \frac{R}{N_0} = \frac{8.317 \text{ joules/mole K}^\circ}{6.023 \times 10^{23} \text{ molecules/mole}}$$

$$= 1.380 \times 10^{-23} \text{ joule/molecule K}^\circ.$$

We shall return to Boltzmann's constant in Chapter 24.

In the last column of Table 23–1 we list calculated values of $\frac{1}{2}Mv_{\mathrm{rms}}{}^2$. As Eq. 23–5 predicts, this quantity (the translational kinetic energy per mole) has (closely) the same value for all gases at the same temperatures, 0° C in this case. From Eq. 23–6 we conclude that at the same temperature $T$ the ratio of the root-mean-square speeds of molecules of two different gases is equal to the square root of the inverse ratio of their masses. That is, from

$$T = \frac{2}{3k}\frac{m_1\overline{v_1{}^2}}{2} = \frac{2}{3k}\frac{m_2\overline{v_2{}^2}}{2}$$

we obtain

$$\sqrt{\frac{\overline{v_1{}^2}}{\overline{v_2{}^2}}} = \frac{v_{1\mathrm{rms}}}{v_{2\mathrm{rms}}} = \sqrt{\frac{m_2}{m_1}}. \tag{23–7}$$

We can apply Eq. 23–7 to the diffusion of two different gases in a container with porous walls placed in an evacuated space. The lighter gas, whose molecules move more rapidly on the average, will escape faster than the heavier one. The ratio of the numbers of molecules of the two gases which find their way through the porous walls for a short time interval is equal to the square root of the inverse ratio of their masses, $\sqrt{m_2/m_1}$. This diffusion process is one method of separating (fissionable) $U^{235}$ (0.7% abundance) from a normal sample of uranium containing mostly (nonfissionable) $U^{238}$ (99.3% abundance). To quote from the Smyth report,†

> As long ago as 1896 Lord Rayleigh showed that a mixture of gases of different atomic weight could be partly separated by allowing some of it to diffuse through a porous barrier into an evacuated space. Because of their higher average speed the molecules of the light gas diffuse through the barrier faster so that the gas which has passed through the barrier (i.e., the "diffusate") is enriched in the lighter constituent and the residual gas which has not passed through the barrier is impoverished in the lighter constituent. The gas most highly enriched in the lighter constituent is the so-called "instantaneous diffusate"; it is the part that diffuses before the impoverishment of the residue has become appreciable. . . . On the assumption that the diffusion rates are inversely proportional to the square roots of the molecular weights‡ the separation factor for the instantaneous diffusate, called the "ideal

---

* See footnote, p. 550.

† *A General Account of the Development of Methods of Using Atomic Energy for Military Purposes* . . . , H. D. Smyth, U. S. Government Printing Office, 1945.

‡ Note that the ratio $m_2/m_1$ of the masses of the two molecules of different gases is the same as the ratio $M_2/M_1$ of their molecular weights, because the molecular weights refer to the same number of molecules. Compare Eq. 23–7.

separation factor" $\alpha$, is given by

$$\alpha = \sqrt{M_2/M_1},$$

where $M_1$ is the molecular weight of the lighter gas and $M_2$ that of the heavier. Applying this formula to the case of uranium will illustrate the magnitude of the separation problem. Since uranium itself is not a gas, some gaseous compound of uranium must be used. The only one obviously suitable is uranium hexafluoride, $UF_6$. . . . Since fluorine has only one isotope, the two important uranium hexafluorides are $U^{235}F_6$ and $U^{238}F_6$; their molecular weights are 349 [gm/mole] and 352 [gm/mole]. Thus if a small fraction of a quantity of uranium hexafluoride is allowed to diffuse through a porous barrier, the diffusate will be enriched in $U^{235}F_6$ by a factor

$$\alpha = \sqrt{\tfrac{352}{349}} = 1.0043 \ldots$$

To separate the uranium isotopes, many successive diffusion stages (i.e., a cascade) must be used. . . . Studies by Cohen and others show that the best flow arrangement for the sucessive stages is that in which half the gas pumped into each stage diffused through the barrier, the other (impoverished) half being returned to the feed of the next lower stage. . . . If one desires to produce 99 per cent pure $U^{235}F_6$, and if one uses a cascade in which each stage has a reasonable overall enrichment factor, then it turns out that roughly 4000 stages are required. . . . Most of the material that eventually emerges from the cascade has been recycled many times. Calculation shows that for an actual uranium-separation plant it may be necessary to force through the barriers of the first stage 100,000 times the volume of gas that comes out the top of the cascade (i.e., as desired product $U^{235}F_6$).

## 23–6 Intermolecular Forces

Forces between molecules are of electromagnetic origin. All molecules contain electric charges in motion. These molecules are electrically neutral in the sense that the negative charge of the electrons is equal and opposite to the charge of the nuclei. This does not mean, however, that molecules do not interact electrically. For example, when two molecules approach each other, the charges on each are disturbed and depart slightly from their usual positions in such a way that the average distance between opposite charges in the two molecules is a little smaller than that between like charges. Hence, an attractive intermolecular force results. This internal rearrangement takes place only when molecules are fairly close together, so that these forces act only over short distances; they are short-range forces. If the molecules come very close together, so that their outer charges begin to overlap, the intermolecular force becomes repulsive. The molecules repel each other because there is no way for a molecule to rearrange itself internally to prevent repulsion of the adjacent external electrons. It is this repulsion on contact that accounts for the billiard-ball character of molecular collisions in gases. If it were not for this repulsion, molecules would move right through each other instead of rebounding on collision.

Let us assume that molecules are approximately spherically symmetrical. Then we can describe intermolecular forces graphically by plotting the mutual potential energy of two molecules, $U$, as a function of distance $r$ between their centers. The force $F$ acting on each molecule is related to the potential energy $U$ by $F = -dU/dr$. In Fig. 23–3$a$ we plot a typical $U(r)$. Here we can imagine one molecule to be fixed at $O$. Then the other molecule will be repelled from $O$ when the slope of $U$ is negative and will be attracted to $O$ when the slope is positive. At $r_0$ no force acts between the molecules; the slope is zero there. In Fig. 23–3$b$ we plot the mutual force $F(r)$ corresponding to this potential energy function. The line $E$ in Fig. 23–3$a$ represents the total mechanical energy of the colliding molecules. The intersection of $U(r)$ with this line is a "turning point" of the motion

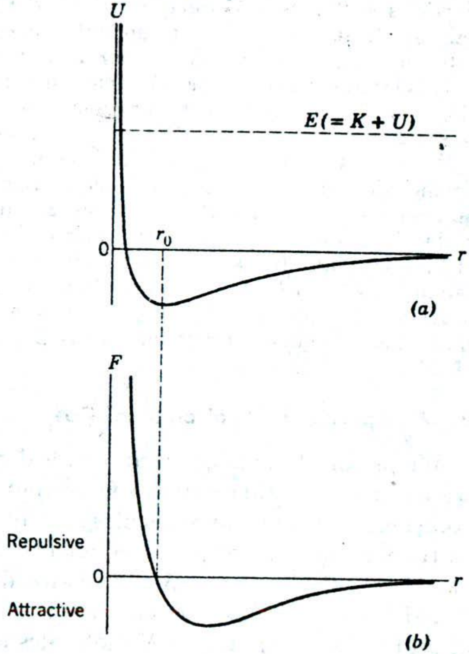


**Fig. 23-3** (*a*) The mutual potential energy of two molecules versus their separation. *E* shows their total mechanical energy ($= K + U$). (*b*) The mutual force, $-dU/dr$, corresponding to this potential energy. *U* is a minimum at $r_0$, at which separation $F = 0$.

(see Section 8-5). The separation of the centers of two molecules at the turning point is the distance of closest approach. The separation distance at which the mutual potential energy is zero may be taken as the approximate distance of closest approach in a collision and hence as the diameter of the molecule. For simple molecules the diameter is about $2.5 \times 10^{-10}$ meter. The forces between molecules practically cease at about $10^{-9}$ meter or 4 diameters apart, so that molecular forces are very short-range ones. The distance $r_0$ at which the potential is a minimum (the equilibrium point) is about $3.5 \times 10^{-10}$ meter for simple molecules. Of course, different molecules have different sizes and internal arrangement of charges so that intermolecular forces vary from one molecule to another. However, they always show the qualitative behavior indicated in the figures.*

In the solid state molecules vibrate about the equilibrium position $r_0$, their total energy *E* being negative, that is, lying below the horizontal axis in Fig. 23-3*a*. The molecules do not have enough energy to escape from the potential valley (that is, from the attractive binding force). The centers of vibration *O* are more or less fixed in a solid. In a liquid the molecules have greater vibrational energy about centers which are free to move but which remain about the same distance from one another. These molecules have their greatest kinetic energy in the gaseous state. In a gas the average distance between the molecules is considerably greater than the effective range of intermolecular forces, and the molecules move in straight lines between collisions. Clerk Maxwell discusses the relation between the kinetic theory model of a gas and the intermolecular forces as follows: "Instead of saying that the particles are hard, spherical, and elastic, we may if we please say that the

---

* See "The Force between Molecules," by B. V. Derjaguin, *Scientific American*, July 1960, for a discussion of the measurement of molecular attractions between macroscopic bodies.

particles are centers of force, of which the action is insensible except at a certain small distance, when it suddenly appears as a repulsive force of very great intensity. It is evident that either assumption will lead to the same results."

It is interesting to compare the measured intermolecular forces with the gravitational force of attraction between molecules. If we choose a separation distance of $4 \times 10^{-10}$ meter, for example, the force between two helium atoms is about $6 \times 10^{-13}$ nt. The gravitational force at that separation is about $7 \times 10^{-42}$ nt, smaller than the intermolecular force by a factor of $10^{29}$! This is a typical result and shows that gravitational forces are negligible in comparison with intermolecular forces. Although the intermolecular forces appear to be small by ordinary standards, we must remember that the mass of a molecule is so small (about $10^{-26}$ kg) that these forces can impart instantaneous accelerations of the order of $10^{15}$ meters/sec² ($10^{14}g$). These accelerations may last for only a very short time, of course, because one molecule can very quickly move out of the range of influence of the other.

## 23–7 Specific Heats of an Ideal Gas

We picture the molecules in an ideal gas as hard elastic spheres; that is, we assume that there are no forces between the molecules except during collisions and that the molecules are not deformed by collisions. If this is so there is no internal potential energy and the internal energy of an ideal gas is entirely kinetic. We have already found that the average translational kinetic energy per molecule is $\frac{3}{2}kT$, so that the internal energy $U$ of an ideal gas containing $N$ molecules is*

$$U = \tfrac{3}{2}NkT = \tfrac{3}{2}\mu RT. \tag{23–8}$$

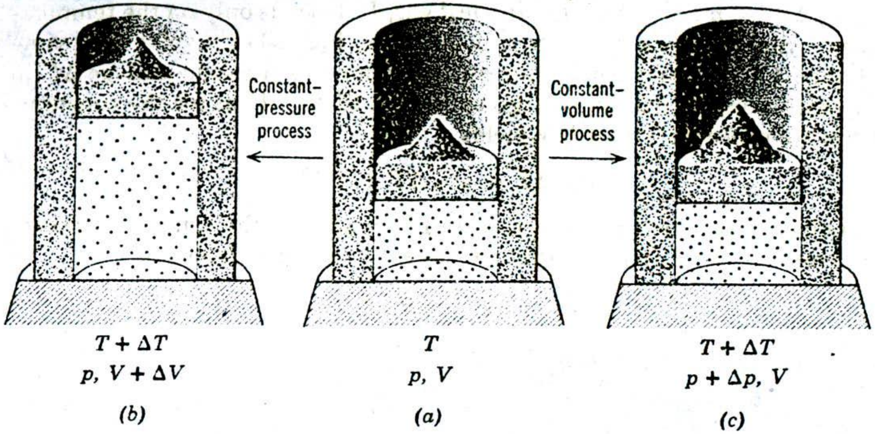
This prediction of kinetic theory says that *the internal energy of an ideal gas is proportional to the Kelvin temperature and depends only on the temperature*, being independent of pressure and volume. With this result we can now obtain information about the specific heats of an ideal gas.

The specific heat of a substance is the heat required per unit mass per unit temperature change. A convenient unit of mass is the mole. The corresponding specific heat is called the molar heat capacity and is represented by $C$. Only two varieties of molar heat capacity are important for gases, namely, that at constant volume, $C_v$, and that at constant pressure, $C_p$.

Let us confine a certain number of moles of an ideal gas in a piston-cylinder arrangement as in Fig. 23–4*a*. The cylinder rests on a heat reservoir whose temperature can be raised or lowered at will, so that we may add heat to the system or remove it, as we wish. The gas has a pressure $p$ such that its upward force on the (frictionless) piston just balances the weight of the piston and its sand load. The state of the system is represented by point $a$ in the $p$-$V$ diagram of Fig. 23–4*d*; this diagram shows two isothermal lines, all points on one corresponding to a

---

* We will see in Section 23–8 that this result applies only to monatomic gases, for which rotational and vibrational energies are not possible. Only in this case can we equate $U$ to the *translational* kinetic energy.



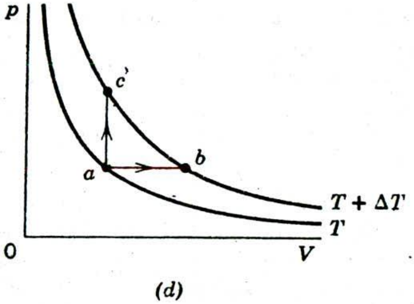


**Fig. 23–4** The temperature of a given mass of gas is raised by the same amount by a constant-pressure process ($a \rightarrow b$) and by a constant-volume process ($a \rightarrow c$).

temperature $T$ and all points on the other to a (higher) temperature $T + \Delta T$.

Now let us raise the temperature of the system by $\Delta T$, by slowly increasing the reservoir temperature. As we do this let us add sand to the piston so that the volume $V$ does not change. This *constant-volume process* carries the system from the initial state of Fig. 23–4*a* to the final state of Fig. 23–4*c*. Equivalently, it goes from point *a* to point *c* in Fig. 23–4*d*. Let us apply the first law of thermodynamics

$$\Delta Q = \Delta U + \Delta W$$

to this process. By definition of $C_v$ we have $\Delta Q = \mu C_v \, \Delta T$. Also, $\Delta W \; (= p \, \Delta V) = 0$ because $\Delta V = 0$. Thus

$$\Delta U = \mu C_v \, \Delta T. \tag{23–9}$$

Let us restore the system to its original state and again raise its temperature by $\Delta T$, this time leaving the sand load undisturbed so that the pressure $p$ does not change. This *constant-pressure process* carries the system from the initial state of Fig. 23–4*a* to the final state of Fig. 23–4*b*. Equivalently, it goes from point *a* to point *b* in Fig. 23–4*d*. Let us apply the first law to *this* process. By definition of $C_p$ we have $\Delta Q = \mu C_p \, \Delta T$.

Also, $\Delta W = p\,\Delta V$. Now for an ideal gas, $U$ depends only on the temperature. Since processes $a \to b$ and $a \to c$ in Fig. 23–4 involve the same change $\Delta T$ in temperature, they must also involve the same change $\Delta U$ in internal energy, namely, that given by Eq. 23–9. Thus for the constant-pressure process the first law yields

$$\mu C_p\,\Delta T = \mu C_v\,\Delta T + p\,\Delta V.$$

Let us apply the equation of state $pV = \mu RT$ to the constant-pressure process $a \to b$. For $p$ constant we have, by taking differences,

$$p\,\Delta V = \mu R\,\Delta T.$$

Combining these equations yields

$$C_p - C_v = R. \tag{23–10}$$

This shows that the molar heat capacity of an ideal gas at constant pressure is always larger than that at constant volume by an amount equal to the universal gas constant $R$ (= 8.31 joules/mole K° or 1.99 cal/mole K°). Although Eq. 23–10 is exact only for an ideal gas, it is nearly true for real gases at moderate pressure (see Table 23–2). Notice that in obtaining this result we did not use the specific relation $U = \frac{3}{2}\mu RT$, but only the fact that $U$ depends on temperature alone.

If we can compute $C_v$, then Eq. 23–10 will give us $C_p$ and vice versa. We *can* obtain $C_v$ by combining Eq. 23–9 with the kinetic theory result for the internal energy of an ideal gas, $U = \frac{3}{2}\mu RT$ (Eq. 23-8). Thus, in the limit of differential changes,

$$C_v = \frac{dU}{\mu\,dT} = \frac{d}{\mu\,dT}\left[\tfrac{3}{2}\mu RT\right] = \tfrac{3}{2}R. \tag{23–11}$$

This result (about 3 cal/mole K°) turns out to be rather good for monatomic gases. However, it is in serious disagreement with values obtained for diatomic and polyatomic gases (see Table 23–2). This suggests that Eq. 23–8 is not generally correct. Since that relation followed directly from the kinetic theory model, we conclude that we must change the model if kinetic theory is to survive as a useful approximation to the behavior of real gases.

▶ **Example 5.** Show that for an ideal gas undergoing an adiabatic process $pV^\gamma$ = a constant, where $\gamma = C_p/C_v$.

Let us apply the first law of thermodynamics

$$\Delta Q = \Delta U + \Delta W.$$

For an adiabatic process $\Delta Q = 0$. For $\Delta W$ we put $p\,\Delta V$. Since the gas is assumed to be ideal, $U$ depends only on temperature and, from Eq. 23–9, $\Delta U = \mu C_v\,\Delta T$. With these substitutions we have

$$0 = \mu C_v\,\Delta T + p\,\Delta V$$

or

$$\Delta T = -\frac{p\,\Delta V}{\mu C_v}.$$

For an ideal gas $pV = \mu RT$, so that, if $p$, $V$, and $T$ are allowed to take on small variations,

$$p\,\Delta V + V\,\Delta p = \mu R\,\Delta T$$

or

$$\Delta T = \frac{p\,\Delta V + V\,\Delta p}{\mu R}.$$

Equating these two expressions and using Eq. 23–10 ($C_p - C_v = R$), we obtain, after some rearrangement,

$$p\,\Delta V C_p + V\,\Delta p C_v = 0$$

Dividing by $pVC_v$ and recalling that, by definition, $C_p/C_v = \gamma$, we get

$$\frac{\Delta p}{p} + \gamma\frac{\Delta V}{V} = 0.$$

In the limiting case of differential changes this reduces to

$$\frac{dp}{p} + \gamma\frac{dV}{V} = 0,$$

which (assuming $\gamma$ to be constant) we can integrate as

$$\ln p + \gamma \ln V = \text{a constant}$$

or

$$pV^{\gamma} = \text{a constant}. \tag{23–12}$$

The value of the constant is proportional to the quantity of gas. In Fig. 23–5 we compare the isothermal and adiabatic behaviors of an ideal gas.

▶ **Example 6.** The compressions and rarefactions in a sound wave are practically adiabatic at audio frequencies. Show that in such a case the speed of sound in an ideal gas is given by

$$v = \sqrt{\frac{\gamma p}{\rho}}.$$

In Chapter 20 we showed the speed of sound to be $v = \sqrt{B/\rho}$, where $\rho$ is the gas density and $B$ is the bulk modulus of the gas, $B = -V(\Delta p/\Delta V)$. However, $B$ will depend on the conditions that prevail as the pressure is changed. If the pressure change is carried out slowly enough so that we can assume the temperature to remain constant we have, in the limit of differential changes,

$$B_{\text{isothermal}} = -V\left(\frac{dp}{dv}\right)_{\text{isothermal}} \tag{23–13}$$

In an isothermal process for an ideal gas we have

$$pV = \text{a constant}$$

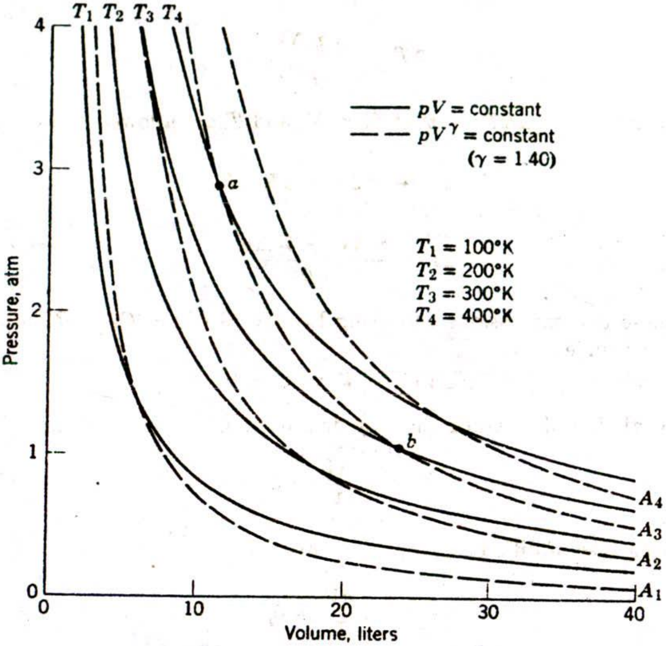


**Fig. 23–5** $T_1$, $T_2$, and $T_3$ show how the pressure of one mole of an ideal gas changes as its volume is changed, the temperature being held constant (isothermal process). $A_1$, $A_2$, and $A_3$ show how the pressure of an ideal gas changes as its volume is changed, no heat being allowed to flow to or from the gas (adiabatic process). An adiabatic *increase* in volume (for example going from $a$ to $b$ along $A_3$) is always accompanied by a *decrease* in temperature, since at $a$, $T$ = 400° K, whereas at $b$, $T$ = 300° K.

or, by differentiation with respect to $V$,

$$p + V\left(\frac{dp}{dV}\right)_{\text{isothermal}} = 0.$$

Combined with Eq. 23–13 this yields

$$B_{\text{isothermal}} = p.$$

In a sound wave, however, the variations are so rapid that the conditions are not isothermal but closely adiabatic. The appropriate bulk modulus is then

$$B_{\text{adiabatic}} = -V\left(\frac{dp}{dV}\right)_{\text{adiabatic}} \qquad (23\text{–}14)$$

In an adiabatic process for an ideal gas we have

$$pV^{\gamma} = \text{a constant}$$

or, by differentiating with respect to $V$,

$$p\gamma V^{\gamma-1} + V^{\gamma}\left(\frac{dp}{dV}\right)_{\text{adiabatic}} = 0.$$

This, combined with Eq. 23–14, yields

$$B_{\text{adiabatic}} = \gamma p$$

and, for the speed of sound,

$$v = \sqrt{\frac{B}{\rho}} = \sqrt{\frac{\gamma p}{\rho}}. \tag{23-15}$$

Newton, in 1710, derived a formula for the speed of sound in which he assumed isothermal rather than adiabatic conditions. He obtained $v = \sqrt{p/\rho}$ rather than the (correct) value of $\sqrt{\gamma p/\rho}$. The error was pointed out by Laplace in 1816. We must remember that, at that date, the concept of energy was not yet understood and the science of thermodynamics did not exist. Does this result modify the result obtained in Example 4? Can you now explain why the speed of sound in a gas is not the same as the root-mean-square speed of the gas molecules? ◀

## 23–8 Equipartition of Energy

A modification of the kinetic theory model designed to explain the specific heats of gases was first suggested by Clausius in 1857. Recall that in our model we assumed a molecule to behave like a hard elastic sphere and we treated its kinetic energy as purely translational. The specific heat prediction was satisfactory for monatomic molecules. Further, because of the success of this simple model in other respects in predicting the correct behavior of gases of all kinds over wide temperature ranges, we feel confident that it is the average kinetic energy of translation which determines what we measure as the temperature of a gas.

However, in the case of specific heats we are concerned with all possible ways of absorbing energy and we must ask whether or not a molecule can store energy internally, that is, in a form other than kinetic energy of translation. This would certainly be so if we pictured a molecule, not as a rigid particle, but as an object with internal structure. For then a molecule could rotate and vibrate as well as move with translational motion. In collisions, the rotational and vibrational modes of motion could be excited, and this would contribute to the internal energy of the gas. Here then is a model which enables us to modify the kinetic theory formula for the internal energy of a gas.

Let us now find the total energy of a system containing a large number of such molecules, where each molecule is thought of as an object having internal structure. The energy will consist of kinetic energy of translation, with terms like $\frac{1}{2}mv_x^2$; of kinetic energy of rotation, with terms like $\frac{1}{2}I\omega_x^2$; of kinetic energy of vibration of the atoms in a molecule, with terms like $\frac{1}{2}\mu v^2$ (where $\mu$ is the reduced mass), and of potential energy of vibration of the atoms in a molecule, with terms like $\frac{1}{2}kx^2$. Although other kinds of energy contributions exist, such as magnetic, for gases we can describe the total energy quite accurately by terms such as these. Although these terms have different origins, they all have the same mathematical form, namely, a positive constant times the square of a quantity which can take on negative or positive values. We can show from statistical mechanics that *when the number of particles is large and Newtonian mechanics holds, all these terms have the same average value, and this average value depends only on the temperature.* In other words, the available energy depends only on the temperature and distributes itself in equal shares to each of the independent ways in which the molecules can absorb energy. This theorem, stated here without proof, is called the *equipartition of energy* and was deduced by Clerk Maxwell. Each such independent mode of energy absorption is called a *degree of freedom.*

From Eq. 23-8 we know that the kinetic energy of translation per mole of gaseous molecules is $\frac{3}{2}RT$. The kinetic energy of translation per mole is the sum of three terms, however, namely $\frac{1}{2}M\overline{v_x^2}$, $\frac{1}{2}M\overline{v_y^2}$, and $\frac{1}{2}M\overline{v_z^2}$. The theorem of equipartition requires that each such term contribute the same amount to the total energy per mole, or $\frac{1}{2}RT$ per degree of freedom.

For *monatomic gases* the molecules have only translational motion (no internal structure in kinetic theory), so that $U = \frac{3}{2}\mu RT$. It follows from Eq. 23-11 that $C_v = \frac{3}{2}R \cong 3$ cal/mole K°. Then from Eq. 23-10, $C_p = \frac{5}{2}R$, and the ratio of specific heat is

$$\gamma = \frac{C_p}{C_v} = \frac{5}{3} = 1.67.$$

For a *diatomic gas* we can think of each molecule as having a dumbbell shape (two spheres joined by a rigid rod). Such a molecule can rotate about any one of three mutually perpendicular axes. However, the rotational inertia about an axis along the rigid rod should be negligible compared to that about axes perpendicular to the rod, so that the rotational energy should consist of only two terms,* such as $\frac{1}{2}I\omega_y^2$ and $\frac{1}{2}I\omega_z^2$. Each rotational degree of freedom is required by equipartition to contribute the same energy as each translational degree, so that, for a diatomic gas having both rotational and translational motion,

$$U = 3\mu(\tfrac{1}{2}RT) + 2\mu(\tfrac{1}{2}RT) = \tfrac{5}{2}\mu RT,$$

or

$$C_v = \frac{dU}{\mu\, dT} = \tfrac{5}{2}R \cong 5 \text{ cal/mole K°}$$

---

* We have already ruled out the possibility that a monatomic molecule could rotate. Actually it could spin about any one of three mutually perpendicular axes if it had any extent, such as a finite sphere. Implicitly, therefore, we have adopted a point mass as our model of the atom. Hence, in a diatomic molecule we are rid of one rotational degree of freedom, for point masses joined by a rigid line have no motion about an axis along that line.

and
$$C_p = C_v + R = \tfrac{7}{2}R,$$
or
$$\gamma = \frac{C_p}{C_v} = \frac{7}{5} = 1.40.$$

For *polyatomic gases*, each molecule contains three or more spheres (atoms) joined together by rods in our model, so that the molecule is capable of rotating energetically about each of three mutually perpendicular axes. Hence, for a polyatomic gas having both rotational and translational motion,

$$U = 3\mu(\tfrac{1}{2}RT) + 3\mu(\tfrac{1}{2}RT) = 3\mu RT,$$
or
$$C_v = \frac{dU}{\mu\, dT} = 3R = 6 \text{ cal/mole K}^\circ,$$
and
$$C_p = 4R,$$
or
$$\gamma = \frac{C_p}{C_v} = 1.33.$$

Let us now turn to experiment to test these ideas. In Table 23-2 we list the experimentally determined molar heat capacities for common gases at 20° C and 1.0 atm. Notice that for monatomic and diatomic gases the values of $C_v$, $C_p$, and $\gamma$ are close to the ideal gas predictions. In some diatomic gases, like chlorine, and in most polyatomic gases the specific heats are larger than the predicted values. Even $\gamma$ shows no simple regularity for polyatomic gases. This suggests that our model is not yet close enough to reality.

We have not yet considered energy contributions from the vibrations of the atoms in diatomic and polyatomic molecules. That is, we can

**Table 23-2**

| Type of Gas | Gas | $C_p$, cal/mole K° | $C_v$, cal/mole K° | $C_p - C_v$ | $\gamma = C_p/C_v$ |
|---|---|---|---|---|---|
| Monatomic | He | 4.97 | 2.98 | 1.99 | 1.67 |
| | A | 4.97 | 2.98 | 1.99 | 1.67 |
| Diatomic | $H_2$ | 6.87 | 4.88 | 1.99 | 1.41 |
| | $O_2$ | 7.03 | 5.03 | 2.00 | 1.40 |
| | $N_2$ | 6.95 | 4.96 | 1.99 | 1.40 |
| | $Cl_2$ | 8.29 | 6.15 | 2.14 | 1.35 |
| Polyatomic | $CO_2$ | 8.83 | 6.80 | 2.03 | 1.30 |
| | $SO_2$ | 9.65 | 7.50 | 2.15 | 1.29 |
| | $C_2H_6$ | 12.35 | 10.30 | 2.05 | 1.20 |
| | $NH_3$ | 8.80 | 6.65 | 2.15 | 1.31 |

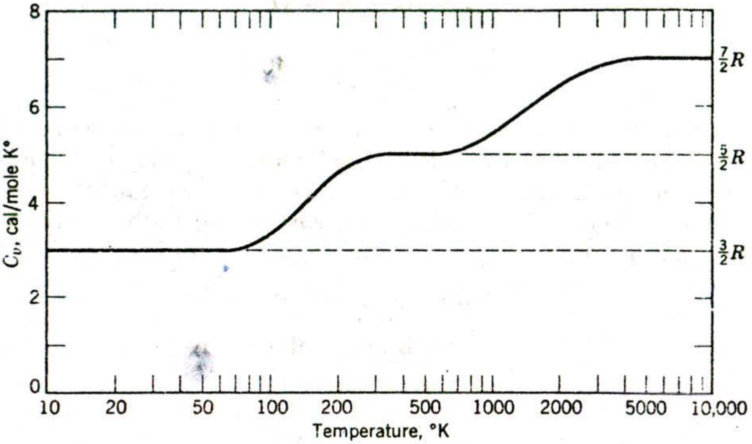


**Fig. 23–6** Variation of the molar heat $C_v$ of hydrogen with temperature. Note that $T$ is drawn on a logarithmic scale.

modify the dumbbell model and join the spheres instead by springs. This new model will greatly improve our results in some cases. However, instead of having a theoretical model for all gases, we now require an empirical model which differs from gas to gas. We can obtain a reasonably good picture of molecular behavior this way and the empirical model is therefore useful; however it ceases to be fundamental.

To see this more clearly, let us consider Fig. 23–6, which shows the variation of the molar heat capacity of hydrogen with temperature. The value of 5 cal/mole K°, which is predicted for diatomic molecules by our model, is characteristic of hydrogen only in the temperature range from about 250 to 750° K. Above 750° K, $C_v$ increases steadily to 7 cal/mole K° and below 250° K, $C_v$ decreases steadily to 3 cal/mole K°. Other gases show similar variations of molar heat with temperature.

Here is a possible explanation. At low temperatures apparently (see Example 7) the hydrogen molecule has translational energy only and, for some reason, cannot rotate. As the temperature rises rotation becomes possible so that at "ordinary" temperatures a hydrogen molecule acts like our dumbbell model. At high temperatures the collisions between molecules cause the atoms in the molecule to vibrate and the molecule ceases to behave as a rigid body. Different gases, because of their different molecular structure, may show these effects at different temperatures. Thus a chlorine molecule appears to vibrate at room temperature.

Although this description is essentially correct, and we have obtained much insight into the behavior of molecules, this behavior contradicts classical kinetic theory. For kinetic theory is based on Newtonian mechanics applied to a large collection of particles, and the equipartition of energy is a necessary consequence of this classical statistical mechanics.

But *if equipartition of energy holds, then, no matter what happens to the total internal energy as the temperature changes, each part of the energy*—translational, rotational, and vibrational—*must share equally in the change.* There is no classical mechanism for changing one mode of mechanical energy at a time in such a system. Kinetic theory requires that the specific heats of gases be independent of the temperature.

Hence, we have come to the limit of validity of classical mechanics when we try to explain the structure of the atom (or molecule). Just as Newtonian principles break down at very high speeds (near the speed of light), so here in the region of very small dimensions they also break down. Relativity theory modifies Newtonian ideas to account for the behavior of physical systems in the region of high speeds. It is quantum physics that modifies Newtonian ideas to account for the behavior of physical systems in the region of small dimensions. Both relativity theory and quantum mechanics are generalizations of classical theory in the sense that they give the (correct) Newtonian results in the regions in which Newtonian physics has accurately described experimental observations. In the following two chapters we shall confine our attention to the very fruitful application of thermodynamics and the kinetic theory to "classical" systems.

► **Example 7.** According to quantum theory the internal energy of an atom (or molecule) is "quantized"; that is, the atom cannot have any of a continuous set of internal energies but *only certain discrete ones.* After being raised from its lowest energy state to some higher one the atom can give up this energy by emitting radiation whose energy equals the difference in energy between the upper and lower internal energy states of the atom.

When two atoms collide, some of their translational kinetic energy may be converted into internal energy of one or both of the atoms. In such a case the collision is inelastic, for translational kinetic energy is not conserved. In a gas, the average translational kinetic energy of an atom is $\frac{3}{2}kT$. When the temperature is raised to a value where $\frac{3}{2}kT$ is about equal to some allowed internal excitation energy of the atom, then an appreciable number of the atoms can absorb enough energy through inelastic collisions to be raised to this higher internal energy state. We can detect this because, after an interval, radiation corresponding to the absorbed energy will be emitted.

(*a*) Compute the average translational kinetic energy per molecule in a gas at room temperature.

We have, for $T = 300°$ K,

$$\begin{aligned}\tfrac{3}{2}kT &= \tfrac{3}{2}(1.38 \times 10^{-23} \text{ joule/molecule °K})(300° \text{ K})\\ &= 6.21 \times 10^{-21} \text{ joule/molecule}\\ &= 3.88 \times 10^{-2} \text{ ev/molecule}\end{aligned}$$

This is about $\frac{1}{25}$ ev per molecule. Some molecules will have larger energies and some smaller energies than this average value.

(*b*) The first allowed (internal) excited state of a hydrogen atom is 10.2 ev above its lowest (ground) state. What temperature is needed to excite a "large" number of hydrogen atoms to emit radiation of this energy?

We require

$$\tfrac{3}{2}kT = 10.2 \text{ ev}$$

and we have from above

$$\tfrac{3}{2}k(300^\circ\text{ K}) = \tfrac{1}{25} \text{ ev}$$

Hence

$$T = 300^\circ\text{ K} \times 10.2/(\tfrac{1}{25}) \simeq 7.5 \times 10^{4\circ}\text{ K}.$$

Actually, because many molecules have energies much greater than the average value, appreciable excitation may occur at somewhat lower temperatures.

We can now appreciate why the kinetic theory assumption, that molecules can be regarded as having no internal structure and collide elastically with one another, holds true at ordinary temperatures. Only at temperatures high enough to give the molecules an average translational kinetic energy comparable to the energy difference between the lowest and the first allowed excited state of the molecule will the internal structure of the molecule change and the collisions become inelastic. Indeed, in retrospect one may say that early evidence that the internal energy of an atom is quantized existed in experiments with gas collisions and that the seeds of quantum theory lay in the kinetic theory of gases.* ◀

## QUESTIONS

**1.** In discussing the fact that it is impossible to apply the laws of mechanics individually to atoms in a macroscopic system, Mayer and Mayer state: "The very complexity of the problem [that is, the fact that the number of atoms is large] is the secret of its solution." Discuss this sentence.

**2.** In kinetic theory we assume that there are a large number of molecules in a gas. Real gases behave like an ideal gas at low densities. Are these statements contradictory? If not, what conclusion can you draw from them?

**3.** We have assumed that the walls of the container are elastic for molecular collisions. Actually, the walls may be inelastic. In practice this makes no difference as long as the walls are at the same temperature as the gas. Explain.

**4.** In large-scale inelastic collisions mechanical energy is lost through internal friction resulting in a rise in temperature owing to increased internal molecular agitation. Is there a loss of mechanical energy to heat in an inelastic collision between molecules?

**5.** What justification is there in neglecting the change in gravitational potential energy of molecules in a gas?

**6.** We have assumed that the force exerted by molecules on the wall of a container is steady in time. How is this justified?

**7.** The average velocity of the molecules in a gas must be zero if the gas as a whole and the container are not in translational motion. Explain how it can be that the average *speed* is not zero.

**8.** By considering quantities which must be conserved in an elastic collision, show that in general molecules of a gas cannot have the same speeds after a collision as they had before. Is it possible, then, for a gas to consist of molecules which all have the same speed?

**9.** Justify the fact that the pressure of a gas depends on the *square* of the speed of its particles by explaining the dependence of pressure on the collision frequency and the momentum transfer of the particles.

---

* See "On Teaching Quantum Phenomena," by Sir N. F. Mott in *Contemporary Physics*, August 1964.

**10.** The gas kinetic temperature in the upper atmosphere (see Eq. 23–5) is of the order of 1000° K. It is quite cold up there. Explain this paradox.

**11.** Why must the time allowed for diffusion separation be relatively short?

**12.** Suppose we want to obtain $U^{238}$ instead of $U^{235}$ as the end product of a diffusion process. Would we use the same process? If not, explain how the separation process would have to be modified.

**13.** Can you describe a centrifugal device for gaseous separation? Is a centrifuge better than a diffusion chamber for separation of gases?

**14.** Would you expect real molecules to be spherically symmetrical? If not, how would the potential energy function of Fig. 23–3 change?

**15.** Explain how we might keep a gas at a constant temperature during a thermodynamic process.

**16.** Explain why the temperature of a gas drops in an adiabatic expansion.

**17.** If hot air rises, why is it cooler at the top of a mountain than near sea level?

**18.** A sealed rubber balloon contains a very light gas. The balloon is released and it rises high into the atmosphere. Describe and explain the thermal and mechanical behavior of the balloon.

**19.** Explain why the specific heat at constant pressure is greater than the specific heat at constant volume.

**20.** It is more common to excite radiation from gaseous atoms by use of electrical discharge than by thermal methods. Why?

## PROBLEMS

**1.** At 0° C and 1.000-atm pressure the densities of air, oxygen, and nitrogen are, respectively, 1.293 kg/meter$^3$, 1.429 kg/meter$^3$, and 1.251 kg/meter$^3$. Calculate the percentage of nitrogen in the air from these data, assuming only these two gases to be present.

**2.** Suppose that, as happened historically, we are given Boyle's law

$$pV = \text{a constant} \quad (\text{constant } T)$$

and Charles' law

$$V/T = \text{a constant} \quad (\text{constant } p)$$

separately. Show how these two laws may be combined to yield

$$pV/T = \text{a constant}.$$

**3.** An air bubble of 20-cm$^3$ volume is at the bottom of a lake 40 meters deep where the temperature is 4° C. The bubble rises to the surface which is at a temperature of 20° C. Take the temperature to be the same as that of the surrounding water and find its volume just before it reaches the surface.

**4.** One mole of an ideal gas undergoes an isothermal expansion. Find the heat flow into the gas in terms of the initial and final volumes and the temperature.

**5.** Calculate the work done in compressing 1.00 mole of oxygen from a volume of 22.4 liters at 0° C and 1.00-atm pressure to 16.8 liters at the same temperature.

**6.** Oxygen gas having a volume of 1.0 liter at 40° C and a pressure of 76 cm-Hg expands until its volume is 1.5 liters and its pressure is 80 cm-Hg. Find the mass in moles of oxygen in the system and its final temperature.

**7.** An automobile tire has a volume of 1000 in.$^3$ and contains air at a gauge pressure of 24 lb/in.$^2$ when the temperature is 0° C. What is the gauge pressure of the air in the tires when its temperature rises to 27° C and its volume increases to 1020 in.$^3$?

**8.** A mercury-filled manometer with two unequal arms is sealed off with the same pressure $p_0$ in the two arms as in Fig. 23-7. The cross-sectional area of the manometer arms is 1.0 cm². With the temperature constant, an additional 10 cm³ of mercury is admitted through the stopcock at the bottom; the level on the left increases 6.0 cm and that on the right increases 4.0 cm. Find the pressure $p_0$.

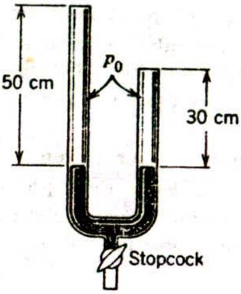


Fig. 23–7

**9.** A thin tube, sealed at both ends, is 1.0 meter long. It lies horizontally, the middle 10 cm containing mercury and the two equal ends containing air at standard atmospheric pressure. If the tube is now turned to a vertical position, by what amount will the mercury be displaced?

**10.** The mass of the $H_2$ molecule is $3.32 \times 10^{-24}$ gm. If $10^{23}$ hydrogen molecules per second strike 2.0 cm² of wall at an angle of 45° with the normal when moving with a speed of $10^5$ cm/sec, what pressure do they exert on the wall?

**11.** (*a*) Determine the average value of the kinetic energy of the molecules of an ideal gas at 0.0° C and at 100° C. (*b*) What is the kinetic energy per mole of an ideal gas at these temperatures?

**12.** (*a*) Compute the root-mean-square speed of an argon atom at room temperature (20° C). (*b*) At what temperature will the root-mean-square speed be half that value? Twice that value?

**13.** (*a*) Compute the temperature at which the root-mean-square speed is equal to the speed of escape from the surface of the earth for hydrogen. For oxygen. (*b*) Do the same for the moon, assuming gravity on its surface to be $0.164g$. (*c*) The temperature high in the earth's upper atmosphere is about 1000° K. Would you expect to find much hydrogen there? Much oxygen?

**14.** At what temperature is the average translational kinetic energy of a molecule equal to the kinetic energy of an electron accelerated from rest through a potential difference of one volt (that is, an energy of 1.0 ev)?

**15.** Show how to find the root-mean-square speeds of helium and argon molecules at 40° C from that of oxygen molecules (460 meters/sec at 0.00° C). The molecular weight of oxygen is 32 gm/mole, of argon 40, of helium 4.

**16.** Compute the number of molecules in a gas contained in a volume of 1.00 cm³ at a pressure of $1.00 \times 10^{-3}$ atm and a temperature of 200° K.

**17.** If the water molecules in 1.0 gm of water were distributed uniformly over the surface of the earth, how many such molecules would there be in 1.0 cm² of the earth's surface?

**18.** Oxygen gas at 273° K and 1.00-atm pressure is confined to a cubical container 10 cm on a side. (*a*) How long does it take a typical molecule to cross the container?

10 cm on a side. Compare the change in gravitational potential energy of an oxygen molecule falling the height of the box with its mean kinetic energy.

**19.** (*a*) Consider an ideal gas at 273° K and 1.0-atm pressure. Imagine that the molecules are for the most part evenly spaced at the centers of identical cubes. Using Avogadro's number and taking the diameter of a molecule to be $3.0 \times 10^{-8}$ cm, find the length of an edge of such a cube and compare this length to the diameter of a molecule. (*b*) Now consider a mole of water having a volume of 18 cm$^3$. Again imagine the molecules to be evenly spaced at the centers of identical cubes. Find the length of an edge of such a cube and compare this length to the diameter of a molecule.

**20.** At 273° F and $1.00 \times 10^{-2}$ atm the density of a gas is $1.24 \times 10^{-5}$ gm/cm$^3$. (*a*) Find $v_{\text{rms}}$ for the gas molecules. (*b*) Find the molecular weight of the gas and identify it.

**21.** *Avogadro's law* states that under the same condition of temperature and pressure equal volumes of gas contain equal numbers of molecules. Derive this law from kinetic theory using Eq. 23–3 and the equipartition of energy assumption.

**22.** *Dalton's law* states that when mixtures of gases having no chemical interaction are present together in a vessel, the pressure exerted by each constituent at a given temperature is the same as it would exert if it alone filled the whole vessel, and that the total pressure is equal to the sum of the partial pressures of each gas. Derive this law from kinetic theory, using Eq. 23–3.

**23.** Plot and interpret physically (*a*) the variation of gas density with temperature for an isobaric (constant-pressure) process and (*b*) the variation of gas density with pressure for an isothermal process.

**24.** Consider a given mass of an ideal gas. Compare curves representing constant-pressure, constant-volume, and isothermal processes on (*a*) a $p$-$V$ diagram, (*b*) a $p$-$T$ diagram and (*c*) a $V$-$T$ diagram. (*d*) How do these curves depend on the mass of gas chosen?

**25.** The mass of a gas molecule can be computed from the specific heat at constant volume. Take $C_v = 0.75$ kcal/kg K° for argon and calculate (*a*) the mass of an argon atom and (*b*) the atomic weight of argon.

**26.** Take the mass of a helium atom to be $6.66 \times 10^{-27}$ kg. Compute the specific heat at constant volume for helium gas.

**27.** Calculate the mechanical equivalent of heat from the value of $R$ and the values of $C_v$ and $\gamma$ for oxygen from Table 23–2.

**28.** The following data are the result of accurate experimental measurements: 1.000 mole of a gas occupies a volume of $2.541 \times 10^{-2}$ meter$^3$ at a pressure of $9.480 \times 10^4$ nt/meter$^2$ when its temperature is 290.0° K. The same mass of gas requires 125.0 cal to raise its temperature from 290.0 to 315.0° K while its volume is held constant. The ratio $(c_p/c_v)$ of its specific heats is 1.430. (*a*) Use these data to compute the mechanical equivalent of heat $J$. (*b*) Account for the fact that your value of $J$ differs from the accepted three-figure value—namely, 4.19 joules/cal.

**29.** One mole of oxygen is heated at a constant pressure starting at 0.00° C. How much heat energy must be added to the gas to double its volume?

**30.** Ten grams of oxygen are heated at constant atmospheric pressure from 27.0 to 127.0° C. How much heat is transferred to the oxygen? What fraction of the heat is used to raise the internal energy of the oxygen?

**31.** How would you explain the observed value of $C_v = 7.50$ cal/mole K° for gaseous $SO_2$ at 15.0° C and 1.00 atm?

**32.** Show that the speed of sound in a gas is independent of the pressure and density.

**33.** Show that the speed of sound in air increases about 2.0 ft/sec for each Celsius degree rise in temperature.

**34.** The speed of sound in different gases at the same temperature depends on the molecular weight of the gas. Show that $v_1/v_2 = \sqrt{M_2/M_1}$ (constant $T$) where $v_1$

is the speed of sound in the gas of molecular weight $M_1$ and $v_2$ is the speed of sound in the gas of molecular weight $M_2$.

**35.** Air at 0.00° C and 1.00-atm pressure has a density of $1.291 \times 10^{-3}$ gm/cm$^3$ and the speed of sound in air is 332 meters/sec at that temperature. Compute the ratio of specific heats of air.

**36.** (*a*) A monatomic ideal gas initially at 17° C is suddenly compressed to one-tenth its original volume. What is its temperature after compression? (*b*) Make the same calculation for a diatomic gas.

**37.** The atomic weight of iodine is 127. A standing wave in a tube filled with iodine gas at 400° K has nodes that are 6.77 cm apart when the frequency is 1000 vib/sec. Is iodine gas monatomic or diatomic?

**38.** A reversible heat engine carries 1.00 mole of an ideal monatomic gas around the cycle shown in Fig. 23–8. Process 1–2 takes place at constant volume, process 2–3 is adiabatic, and process 3–1 takes place at a constant pressure. (*a*) Compute the approximate numerical values for the heat $\Delta Q$, the change in internal energy $\Delta U$, and the work done $\Delta W$, for each of the three processes and for the cycle as a whole. (*b*) If the initial pressure at point 1 is 1.00 atm, find the pressure and the volume at points 2 and 3.

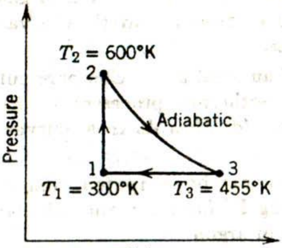


Fig. 23–8

**39.** A mass of gas occupies a volume of 4.0 liters at a pressure of 1.0 atm and a temperature of 300° K. It is compressed adiabatically to a volume of 1.0 liter. Determine (*a*) the final pressure and (*b*) the final temperature, assuming it to be an ideal gas for which $\gamma = 1.5$.

**40.** An ideal gas expands adiabatically from an initial temperature $T_1$ to a final temperature $T_2$. Prove that the work done by the gas is $C_v(T_1 - T_2)$.

**41.** (*a*) A liter of gas with $\gamma = 1.3$ is at 273° K and 1.0-atm pressure. It is suddenly compressed to half its original volume. Find its final pressure and temperature. (*b*) The gas is now cooled back to 0° C at constant pressure. What is its final volume?

**42.** (*a*) Show that the variation in pressure in the earth's atmosphere, assumed to be isothermal, is given by $p = p_0 e^{-Mgy/RT}$ where $M$ is the molecular weight of the gas. (See Example 1, Chapter 17.) (*b*) Show also that $n = n_0 e^{-Mgy/RT}$ where $n$ is the number of molecules per unit volume.

**43.** A hydrogen atom, in its lowest (ground) state and moving with 13-ev kinetic energy, collides head-on with another hydrogen atom which is *at rest* in its ground state. (*a*) Use the conservation laws of energy and momentum to show that this collision must be elastic. The first allowed excited state is about 10.2 ev above the ground state. (*b*) Show that the minimum initial kinetic energy that the incident atom needs to raise one of the atoms to the first excited state is *twice* the energy difference between ground state and first excited state.

# Kinetic Theory of Gases—II

## CHAPTER 24

### 24–1 Mean Free Path

Between successive collisions a molecule in a gas moves with constant speed along a straight line. The average distance between such successive collisions is called the *mean free path* (Fig. 24–1). If molecules were points, they would not collide at all and the mean free path would be infinite. Molecules, however, are not points and hence collisions occur. If they were so numerous that they completely filled the space available to them, leaving no room for translational motion, the mean free path would be zero. Thus the mean free path is related to the size of the molecules and to their number per unit volume.

Consider the molecules of a gas to be spheres of diameter $d$. The cross section for a collision is then $\pi d^2$. That is, a collision will take place when the centers of two molecules approach within a distance $d$ of one another. An equivalent description of collisions

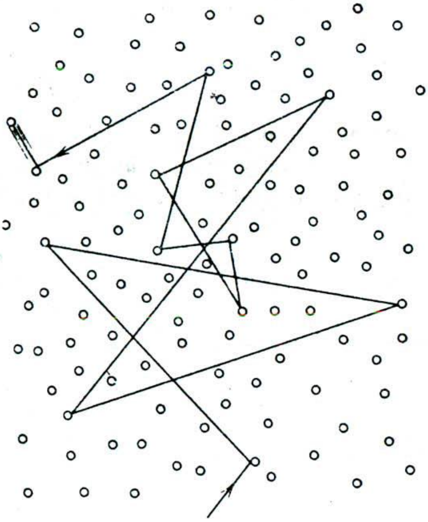


**Fig. 24–1** A molecule traveling through a gas, colliding with other molecules in its path. Of course, all the other molecules are moving in a similar fashion.

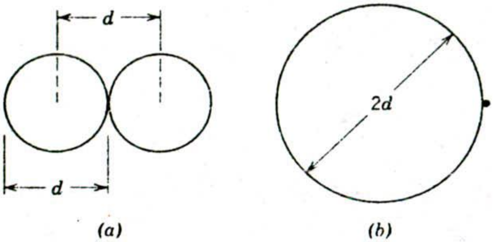


**Fig. 24–2** (*a*) If a collision occurs when two molecules come within a distance $d$ of each other, the process can be treated equivalently (*b*) by thinking of one molecule as having an effective diameter $2d$ and the other as being a point mass.

made by any one molecule is to regard that molecule as having a diameter $2d$ and all other molecules as point particles (see Fig. 24–2).

Imagine a typical molecule of equivalent diameter $2d$ moving with speed $v$ through a gas of equivalent point particles and let us assume, for the time being, that the molecule and the point particles exert no forces on each other. In time $t$ our molecule will sweep out a cylinder of cross-sectional area $\pi d^2$ and length $vt$. If $n$ is the number of molecules per unit volume the cylinder will contain $(\pi\, d^2 vt)\, n$ particles (see Fig. 24–3). Since our molecule and the point particles *do* exert forces on each other, this will be the number of collisions experienced by the molecule in time $t$. The cylinder of Fig. 24–3 will, in fact, be a broken one, changing direction with every collision.

The mean free path $\bar{l}$ is the average distance between successive collisions. Hence, $\bar{l}$ is the total distance $vt$ covered in time $t$ divided by the number of collisions that take place in this time, or

$$\bar{l} = \frac{vt}{\pi\, d^2 nvt} = \frac{1}{\pi n d^2}.$$

This equation is based on the picture of a molecule hitting stationary targets. Actually the molecule hits moving targets. The collision frequency is increased as a result (see below) and the mean free path is reduced to

$$\bar{l} = \frac{1}{\pi \sqrt{2}\, n d^2}. \tag{24–1}$$

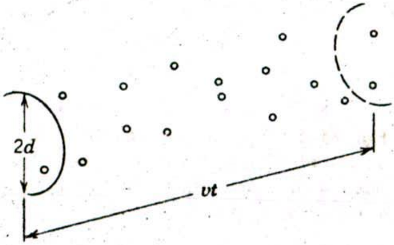


**Fig. 24–3** A molecule of equivalent diameter $2d$ traveling with speed $v$ sweeps out a cylinder of base $\pi d^2$ and length $vt$ in a time $t$. It suffers a collision with every other molecule whose center lies within this cylinder.

When the target molecules are moving, the two $v$'s in the first equation above are not the same. The one in the numerator ($= \bar{v}$) is the mean molecular speed measured with respect to the container. The one in the denominator ($= \bar{v}_{\text{rel}}$) is the mean *relative* speed with respect to other molecules; it is this relative speed that determines the collision rate.

We can see qualitatively that $\bar{v}_{\text{rel}} > \bar{v}$. Thus two molecules of speed $v$ moving toward each other have a relative speed of $2v$ ($> v$); two molecules with speed $v$ moving at right angles on a collision course have a relative speed of $\sqrt{2}\,v$ (also $> v$); two molecules moving with speed $v$ in the same direction have a relative speed of zero ($< v$). Thus molecules arriving from *all of the forward hemisphere* and *part of the backward hemisphere* have $\bar{v}_{\text{rel}} > \bar{v}$. The molecules arriving from the rest of the backward hemisphere have $\bar{v}_{\text{rel}} < \bar{v}$ but, since their numbers are smaller, they do not determine the nature of the average over both hemispheres, which yields $\bar{v}_{\text{rel}} > \bar{v}$. A quantitative calculation, taking into account the actual speed distribution of the molecules, gives $\bar{v}_{\text{rel}} = \sqrt{2}\,\bar{v}$.

▶ **Example 1.** Let us calculate the magnitude of the mean free path and the collision frequency for air molecules at 0° C and 1-atm pressure.

We take $2 \times 10^{-8}$ cm as an effective molecular diameter $d$. For the conditions stated, the average speed of air molecules is about $1 \times 10^5$ cm/sec and there are about $3 \times 10^{19}$ molecules/cm$^3$. The mean free path is then

$$\bar{l} = \frac{1}{\pi\sqrt{2}\,nd^2} = \frac{1}{\pi\sqrt{2}\,(3 \times 10^{19}/\text{cm}^3)(2 \times 10^{-8}\text{ cm})^2}$$

$$= 2 \times 10^{-5}\text{ cm}.$$

This is about a thousand molecular diameters.

The corresponding collision frequency is

$$\frac{\bar{v}}{\bar{l}} = (1 \times 10^5\text{ cm/sec})/(2 \times 10^{-5}\text{ cm})$$

$$= 5 \times 10^9/\text{sec}.$$

Thus, on the average, *each molecule* makes five billion collisions per second! ◀

In the earth's atmosphere the mean free path of air molecules at sea level (760 mm-Hg) is about $10^{-5}$ cm. At 100 km above the earth ($10^{-3}$ mm-Hg) the mean free path is 1 meter. At 300 km ($10^{-6}$ mm-Hg) it is 10 km or 6 miles, and yet there are about $10^8$ molecules/cm$^3$ in this region. This emphasizes that molecules are indeed small. At great enough heights the mean free path concept fails because the upward-directed molecules follow ballistic paths and may escape from the atmosphere.

In the laboratory the mean free path concept is useful in situations such as that of Example 1. In even modest laboratory vacuums, however, it loses some of its meaning because nearly all the collisions are with the wall of the containing vessel rather than with other molecules. Consider a box 10 cm on edge containing air at $10^{-6}$ mm-Hg pressure. The mean free path (see above) is 6 miles, so that collisions between molecules are rare indeed. And yet this box contains about $10^{12}$ molecules!

Even in a finite "box," however, there are some conditions in which particles can travel great distances without striking the walls. In a typical proton synchrotron, used to accelerate protons to the billion-electron-volt range of energies, the protons are constrained by a magnetic field to move in a circular path and may travel *several hundred thousand miles* during the acceleration process. Mean free path considerations are important if the accelerating protons are to have essentially no collisions with residual air molecules. In this case the effective cross section of the proton is so much smaller than that of the air molecules that if we have a vacuum of about $10^{-6}$ mm-Hg, there is essentially no beam loss by proton scattering from gas molecules inside the vacuum chamber.

## 24–2 Distribution of Molecular Speeds

In Chapter 23 we discussed the root-mean-square speed of the molecules of a gas. However, the speeds of individual molecules vary over a wide range of magnitude; there is a characteristic distribution of molecular speeds for a given gas which depends, as we will see below, on the temperature. If all the molecules of a gas had the same speed $v$, this situation would not persist for very long because the molecular speeds would be changed by collisions. However, we do not expect many molecules to have speeds $\ll v_{\text{rms}}$ (that is, near zero) or $\gg v_{\text{rms}}$ because such extreme speeds would require an unlikely sequence of preferential collisions.

Clerk Maxwell first solved the problem of the most probable distribution of speeds in a large number of molecules of a gas. His molecular speed distribution law, for a sample of gas containing $N$ molecules, is*

$$N(v) = 4\pi N(m/2\pi kT)^{3/2} v^2 e^{-mv^2/2kT} \tag{24–2}$$

In this equation $N(v)\,dv$ is the number of molecules in the gas sample having speeds between $v$ and $v + dv$. $T$ is the absolute temperature, $k$ is Boltzmann's constant, and $m$ is the mass of a molecule. Note that for a given gas the speed distribution depends only on the temperature. We find $N$, the total number of molecules in the sample, by adding up (that is, by integrating) the number present in each differential speed interval from zero to infinity, or

$$N = \int_0^\infty N(v)\,dv. \tag{24–3}$$

The unit of $N(v)$ is, say, molecules/(cm/sec).

In Fig. 24–4 we plot the Maxwell distribution of speeds for molecules of oxygen at two different temperatures. The number of molecules having a speed between $v_1$ and $v_2$ equals the area under the curve between the vertical lines at $v_1$ and $v_2$. As Eq. 24–3 shows, the area under the speed distribution curve, which is the integral in that equation, is equal to the total number of molecules in the sample. At any temperature the number of

---

* A derivation of Eq. 24–2 appears in Supplementary Topic IV.

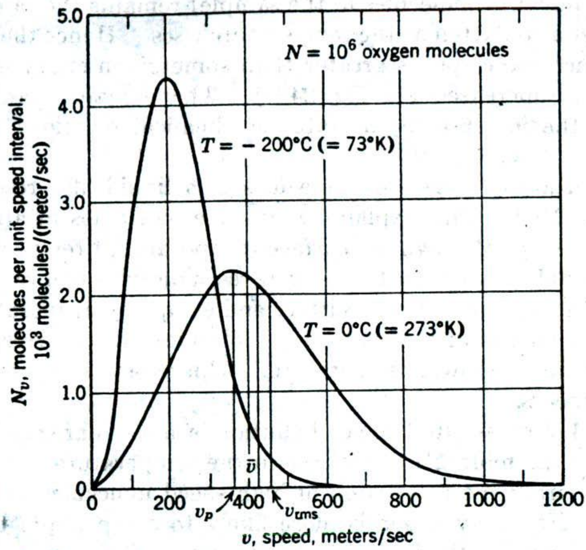


**Fig. 24–4** The Maxwellian distribution of speeds of $10^6$ oxygen molecules at two different temperatures. The number of molecules within a certain range of speeds (say, 300 to 600 meters/sec) is the area under this section of the curve. The complete area under either curve is the total number of molecules (equals $10^6$); this area must be the same for each temperature if, as in this case, the curves refer to a given number of molecules.

molecules in a given speed interval* $\Delta v$ increases as the speed increases up to a maximum (the most probable speed $v_p$) and then decreases asymptotically toward zero. The distribution curve is not symmetrical about the most probable speed because the lowest speed must be zero, whereas there is no classical limit to the upper speed a molecule can attain. In this case the *average speed* $\bar{v}$ is somewhat larger than the most probable value. The root-mean-square value, $v_{rms}$, being the square root of the sum of the *squares* of the speeds, is still larger.

As the temperature increases, the root-mean-square speed $v_{rms}$ (and $\bar{v}$ and $v_p$ as well) increases, in accord with our microscopic interpretation of temperature. The range of typical speeds is now greater, so that the distribution broadens. Since the area under the distribution curve (which

---

* We cannot simply plot the "number of particles having speed $v$" against $v$, because there are a finite number of particles and an infinite number of possible speeds. Hence, the probability that a particle has a precisely stated speed, such as 279.343267 · · · meters/sec, is exactly zero. However, we can divide the range of speeds into intervals and the probability that a particle has a speed somewhere in a given interval (such as 279 meters/sec to 280 meters/sec) has a definite nonzero value.

is the total number of molecules in the sample) remains the same, the distribution must also flatten as the temperature rises. Hence the number of molecules which have speeds greater than some given speed increases as the temperature increases (see Fig. 24–4). This explains many phenomena, such as the increase in the rates of chemical reactions with rising temperature.

The distribution of speeds of molecules in a liquid also resembles the curves of Fig. 24–4. This explains why some molecules in a liquid (the fast ones) can escape through the surface (evaporate) at temperatures well below the normal boiling point. Only these molecules can overcome the attraction of the molecules in the surface and escape by evaporation. The average kinetic energy of the remaining molecules drops correspondingly, leaving the liquid at a lower temperature. This explains why evaporation is a cooling process.

From Eq. 24–2 we see that the distribution of molecular speeds depends on the mass of the molecule as well as on the temperature. The smaller the mass, the larger the proportion of high-speed molecules at any given temperature. Hence hydrogen is more likely to escape from the atmosphere at high altitudes than oxygen or nitrogen. The moon may have a tenuous atmosphere. For the molecules in this atmosphere not to have a great probability of escaping from the weak gravitational pull of the moon, even at the low temperatures there, we would expect them to be molecules or atoms of the heavier elements. Evidence points to the heavy inert gases, such as krypton and xenon, which were produced largely by radioactive decay early in the moon's history. The atmospheric pressure on the moon is believed to be about $10^{-13}$ of the earth's atmospheric pressure.

▶ **Example 2.** The speeds of ten particles in meters/sec are 0, 1.0, 2.0, 3.0, 3.0, 3.0, 4.0, 4.0, 5.0, and 6.0. Find (*a*) the average speed, (*b*) the root-mean-square speed, and (*c*) the most probable speed of these particles.

(*a*) The average speed is

$$\bar{v} = \frac{0 + 1.0 + 2.0 + 3.0 + 3.0 + 3.0 + 4.0 + 4.0 + 5.0 + 6.0}{10} = 3.1 \text{ meters/sec.}$$

(*b*) The mean-square speed is

$$\overline{v^2} = \frac{0 + (1.0)^2 + (2.0)^2 + (3.0)^2 + (3.0)^2 + (3.0)^2 + (4.0)^2 + (4.0)^2 + (5.0)^2 + (6.0)^2}{10}$$

$$= 12.5 \text{ meters}^2/\text{sec}^2$$

and the root-mean-square speed is

$$v_{\text{rms}} = \sqrt{12.5 \text{ meters}^2/\text{sec}^2} = 3.5 \text{ meters/sec.}$$

(*c*) Of the ten particles three have speeds of 3.0 meters/sec, two have speeds of 4.0 meters/sec, and the other five each have a different speed. Hence, the most probable speed of a particle $v_p$ is

$$v_p = 3.0 \text{ meters/sec.}$$

**Example 3.** Use Eq. 24-2 to determine the average speed $\bar{v}$, the root-mean-square speed $v_{\mathrm{rms}}$, and the most probable speed $v_p$ of the molecules in a gas in terms of the gas parameters.

The quantity $N(v)\,dv$ is the number of particles in the sample having a speed between $v$ and $v + dv$, $N(v)$ being given by Eq. 24-2. We find the average speed $\bar{v}$ in the usual way: we multiply the number of particles in each speed interval by a speed $v$ characteristic of that interval; we sum these products over all speed intervals and we divide by the total number of particles. Replacing the summation by an integration, we obtain

$$\bar{v} = \frac{\int_0^\infty N(v)v\,dv}{N}$$

Substituting Eq. 24-2 for $N(v)$ and integrating* we obtain

$$\bar{v} = \sqrt{\frac{8}{\pi}\frac{kT}{m}} = 1.59\sqrt{\frac{kT}{m}} \quad \text{(average speed)}.$$

The mean-square speed is given by

$$\overline{v^2} = \frac{\int_0^\infty N(v)v^2\,dv}{N}$$

which yields

$$v_{\mathrm{rms}} = \sqrt{\overline{v^2}} = \sqrt{\frac{3kT}{m}} = 1.73\sqrt{\frac{kT}{m}} \quad \text{(root-mean-square speed)}.$$

The most probable speed $v_p$ is the speed at which $N(v)$ has its maximum value. It is given by requiring that

$$\frac{dN(v)}{dv} = 0.$$

By substituting Eq. 24-2 for $N(v)$ we obtain, as the student should show,

$$\bar{v}_p = \sqrt{\frac{2kT}{m}} = 1.41\sqrt{\frac{kT}{m}} \quad \text{(most probable speed)}.$$

In Fig. 24-4 we show $v_p$, $\bar{v}$, and $v_{\mathrm{rms}}$ at 0° C for a molecular speed distribution in oxygen. ◀

## 24-3 Experimental Confirmation of the Maxwellian Distribution

Maxwell derived his distribution law for molecular speeds (Eq. 24-2) in 1859. At that early date it was not possible to check this law by direct measurement and, indeed, it was not until 1920 that Stern made the first serious attempt to do so. Techniques improved rapidly in the hands of various workers but it was not until 1955 that a high-precision experimental verification of the law (for gas molecules) was provided, by Miller and Kusch.

Their apparatus is shown in Fig. 24-5. The walls of oven $O$ were heated, in one

---

* Let $\lambda = m/2kT$. From tables of integrals,

$$\int_0^\infty v^2 e^{-\lambda v^2}\,dv = \frac{1}{4}\sqrt{\frac{\pi}{\lambda^3}}; \quad \int_0^\infty v^3 e^{-\lambda v^2}\,dv = \frac{1}{2\lambda^2}; \quad \int_0^\infty v^4 e^{-\lambda v^2}\,dv = \frac{3}{8}\sqrt{\frac{\pi}{\lambda^5}}.$$

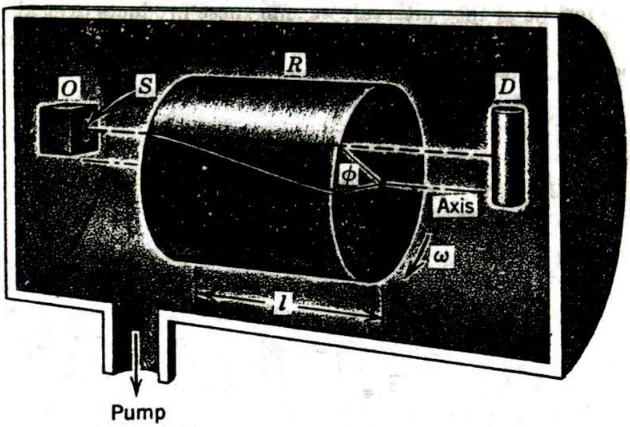


**Fig. 24–5** The apparatus used by Miller and Kusch to verify the Maxwell speed distribution law. The mechanism for rotating the cylinder is not shown. The whole apparatus is highly evacuated to reduce collisions with the residual gas molecules of the thallium molecules in the beam emerging from slit $S$.

set of experiments, to a uniform temperature of 870 ± 4° K, some thallium having been placed in the oven. At this temperature thallium vapor, at a pressure of $3.2 \times 10^{-3}$ mm-Hg, fills the oven. Some molecules of thallium vapor escape from slit $S$ into the highly evacuated space outside the oven, falling on the rotating cylinder $R$. This cylinder, of length $l$, has a number of helical grooves cut into it, only one of them being shown in Fig. 24–5. For a given angular speed $\omega$ of the cylinder, only molecules of a sharply defined speed $v$ can pass along the grooves without striking the walls. The speed $v$ can be found from:

$$\text{time of travel along the groove} = \frac{l}{v} = \frac{\phi}{\omega},$$

or

$$v = l\omega/\phi \tag{24–4}$$

in which $\phi$ (see Fig. 24–5) is the angular displacement between the entrance and the exit of a helical groove. Thus the rotating cylinder is a *velocity selector*, the speed selected being proportional to the (controllable) angular speed $\omega$, as Eq. 24–4 shows. One observes the beam intensity recorded by detector $D$ as a function of the selected speed $v$. Figure 24–6 shows the remarkable agreement between theory (the solid line) and experiment (the triangles and circles) for thallium vapor.

The distribution of speeds in the *beam* (as distinguished from the distribution of speeds in the *oven*) is not proportional to $v^2e^{-mv^2/2kt}$, as in Eq. 24–2, but to $v^3e^{-mv^2/2kT}$. Consider a group of molecules in the oven whose speeds lie within a certain small range $v_1$ to $v_1 + \Delta v$, where $v_1$ is less than the most probable speed $v_p$ We can always find another equal speed interval $\Delta v$, extending from $v_2$ to $v_2 + \Delta v$ where $v_2$, which will be greater than $v_p$, is chosen so that the two speed intervals contain the same number of molecules. However, more molecules in the latter interval than in the former will escape from slit $S$ to form the beam, because molecules in the latter interval "bombard" the slit with a greater frequency, by precisely the factor $v_2/v_1$. Thus, other things being equal, fast molecules are favored in escaping from the oven, just in proportion to their speeds, and the molecules in the beam have a "$v^3$" rather than a "$v^2$" distribution. This effect is allowed for in computing the theoretical curve of Fig. 24–6.

Rainwater and Havens (1946) also provided a convincing experimental check of the Maxwell speed distribution law by using a "gas" of neutrons. The neutrons were produced (as fast neutrons) in continuous series of short bursts in a cyclotron and allowed to fall on a block of paraffin. By repeated collisions with the nuclei of the block, the neutrons rapidly slowed down and came into thermal equilibrium with the block, behaving like a "neutron gas" in a container. The container, however, is a leaky one because neutrons diffuse out through the walls of the block and move across the laboratory. It is possible, by electronic means, to measure the time between the production of the neutrons in the cyclotron and their arrival at a distant detector after escaping from the paraffin block. Thus one can measure the speed distribution in a collimated beam of escaping neutrons and can compare it to the prediction of Maxwell; the agreement of theory and experiment is excellent.

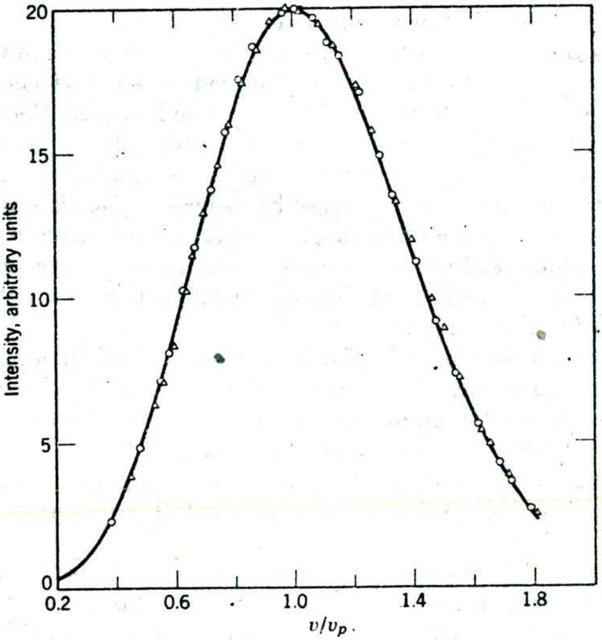


**Fig. 24–6** The solid line shows Maxwell's molecular speed distribution. The circles (○) are experimental points for thallium atoms emerging from an oven at 870° K; the triangles (△) correspond to 944° K. The horizontal scale is a plot of $v/v_p$ where $v_p$ is the most probable speed. When speeds are plotted in this way the distributions for different temperatures should fall on the same curve. At 870° K, $v_p = 376$ meters/sec and at 944° K, it is 395 meters/sec. From Miller and Kusch, *Physical Review*, 99, 1314 (1955).

Although the Maxwell speed distribution for gases agrees remarkably well with observations under ordinary conditions, it fails at high densities, where the basic assumptions of the classical kinetic theory fail. In these regions we must use speed distributions founded on the principles of quantum physics, the Fermi-Dirac and the Bose-Einstein distributions. These quantum distributions agree closely with the Maxwell distribution in the classical region (low density) and agree with experiment where the classical distribution fails. Hence there are limits to the applicability of the Maxwell distribution, as in fact there are to any theory.

## 24–4 Brownian Motion

The prominence given to atomic and molecular theory during the last quarter of the nineteenth century was deplored by many able scientists. In spite of the many quantitative agreements between kinetic theory and the behavior of gases, no proof of the separate existence of atoms and molecules had been obtained, nor had any observation been made that could really demonstrate the continuous motions of the molecules. Ernst Mach (1838–1916) saw no point to "thinking of the world as a mosaic, since we cannot examine its individual pieces of stone." It had been established rather early in the development of kinetic theory that an atom should be about $10^{-7}$ cm or $10^{-8}$ cm in diameter. No one actually expected to see an atom or detect the effect of a single atom.

The leader of the opposition to the atomic theory was Wilhelm Ostwald. He was a champion of the principle of the conservation of energy and regarded energy as the ultimate reality. Ostwald argued that with a thermodynamical treatment of a process we know all that is essential about the process and that further mechanical assumptions about the mechanism of the reactions are unproved hypotheses. He abandoned the atomic and molecular theories and fought to free science "from hypothetical conception which lead to no immediate experimentally verifiable conclusions." Other prominent physicists were reluctant to admit the atom as an established scientific fact.

Ludwig Boltzmann felt compelled to protest this attitude in an article in 1897, stressing the indispensability of atomism in natural science. The progress of science is often guided by the analogies of nature's processes which occur in the minds of investigators. Kinetic theory was such a mechanical analogy. As with most analogies it suggests experiments to test the validity of our mental pictures and leads to further investigations and clearer knowledge.

As is always true in such controversies in science, the decision rests with experiment. The earliest and most direct experimental evidence for the reality of atoms was the proof of the atomic kinetic theory provided by the quantitative studies of Brownian motion. These observations convinced both Mach and Ostwald of the validity of the kinetic theory and the atomic description of matter on which it rests. The atomic theory gained unquestioned acceptance in later years when a wide variety of experiments all led to the same values of the fundamental atomic constants.

Brownian motion is named after the English botanist Robert Brown who discovered in 1827 that pollen suspended in water shows a continuous random motion when viewed under a microscope. At first these motions were considered a form of life, but it was soon found that small inorganic particles behave similarly. There was no quantitative explanation of this phenomenon until the development of kinetic theory. Then, in 1905, Albert Einstein developed a theory of Brownian motion.* In his *Autobiographical Notes*, Einstein writes, "My major aim in this was to find facts which would guarantee as much as possible the existence of atoms of definite size. In the midst of this I discovered that, according to atomistic theory, there would have to be a movement of suspended microscopic particles open to observation, without knowing that observations concerning the Brownian motion were already long familiar."

The basic assumption made by Einstein was that particles suspended in a liquid or a gas share in the thermal motions of the medium and that on the average the translational kinetic energy of each particle is $\frac{3}{2}kT$, in accordance with the principle of equipartition of energy. In this view the Brownian motions result from impacts by molecules of the fluid, and the suspended particles acquire the same mean kinetic energy as the molecules of the fluid.

The suspended particles are extremely large compared to the molecules of the fluid and are being continually bombarded on all sides by them. If the particles are sufficiently large and the number of molecules is sufficiently great, equal numbers of molecules strike the particles on all sides at each instant. For smaller particles and fewer molecules the number of molecules striking various sides of the particle at any instant, being merely a matter of chance, may not be equal; that is, fluctuations occur. Hence the particle at each instant suffers an unbalanced force causing it to move this way or that. The particles therefore act just like very large molecules in the fluid, and their motions should be qualitatively the same as the motions of the fluid molecules. If Avogadro's number were infinite there would be no statistical unbalance (fluctuations) and no Brownian motion. If Avogadro's number were very small, the Brownian motion would be very large. Hence we should be able to deduce the value of Avogadro's number from observations of the Brownian motion. Deeply ingrained in this picture is the idea of molecular motion and the smallness of molecules. The Brownian motion therefore offers a striking experimental test of the kinetic theory hypotheses.

The suspended particles are under the influence of gravity and would settle to the bottom of the fluid were it not for the molecular bombardment opposing this tendency. Since the suspended particles behave like gas molecules we are not surprised to learn that, as for molecules in the atmosphere, their density drops off exponentially with respect to height in the

---

* Einstein's theory appeared as an article in the same issue of the *Annalen der Physik* which contained his famous paper on the theory of relativity and also his paper on the theory of the photoelectric effect. It was for his work on the photoelectric effect that he won the Nobel prize in 1921.

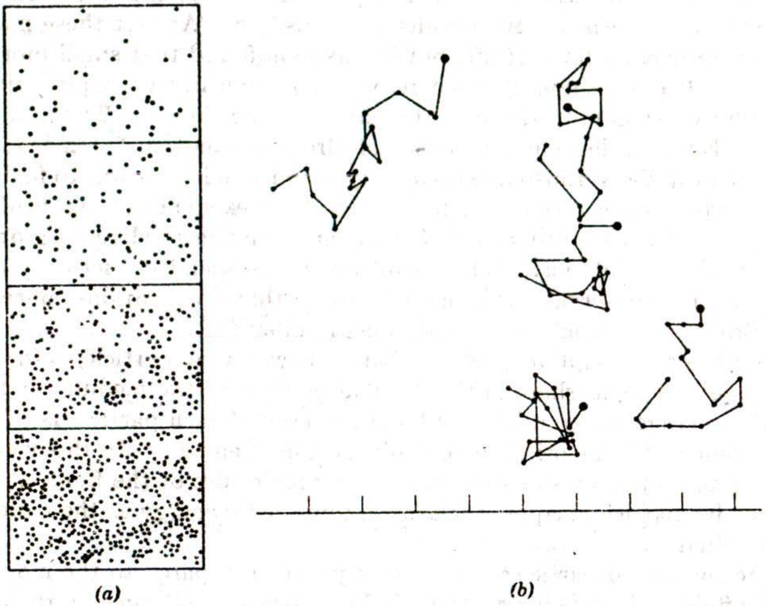


**Fig. 24–7** (*a*) A gum resin suspension contained in a glass vessel viewed in a microscope by Perrin in 1909. At first the distribution of particles was uniform, but in time they settled to the distribution shown. The particles have a diameter of $0.6 \times 10^{-3}$ cm and the horizontal lines are $10 \times 10^{-3}$ cm apart. (*b*) Sketch by V. Henri in 1908 from his cinematographic study of Brownian movement. Henri used a microscope with a motion-picture camera which ran 20 frames/sec, each exposure being $\frac{1}{320}$ sec. The zigzag lines show the position of five rubber particles as recorded by successive frames. The lines do not represent the actual paths of the particles for between exposures the particles may have traveled a similar erratic path. The scale at the bottom is divided into microns ($10^{-4}$ cm).

fluid; they from a "miniature atmosphere"; see Example 1, Chapter 17; Problem 42, Chapter 23; and Problem 15, this chapter. Jean Perrin, a French physical chemist, confirmed this prediction in 1908 by determining the numbers of small particles of gum resin suspended at different heights in a liquid drop (Fig. 24–7, left). From his data he deduced a value of Avogadro's number $N_0 = 6 \times 10^{23}$ particles/mole. Perrin also made measurements of the displacements of Brownian particles during many equal time intervals and found that they have the statistical distribution required by kinetic theory and the root-mean-square displacement predicted by Einstein (Fig. 24–7, right).

Among the many subsequent experiments was that of Kappler, in 1931, who observed the Brownian motion of a rather gross object, a small mirror (area 0.7 mm$^2$), mounted on a fine torsion fiber with light reflected from the mirror to a moving photographic film. The mirror is mounted in a chamber with gas at low pressure ($10^{-2}$ mm-Hg); the record on the moving film yields the function $\theta(t)$

(angular displacement as a function of time). This shows clearly the rotational Brownian motion of the mirror which consists of a series of angular displacements produced by unbalanced impacts from the molecules. As the gas pressure is lowered, there is a gradual decrease in the motion. From the photographic record we can find the angular displacement $\theta$ and the angular velocity $\omega$. The equipartition of energy principle requires that

$$\tfrac{1}{2}I\overline{\omega^2} = \tfrac{1}{2}\kappa\overline{\theta^2} = \tfrac{1}{2}kT,$$

for $\frac{1}{2}I\overline{\omega^2}$ is the average rotational kinetic energy of the system and $\frac{1}{2}\kappa\overline{\theta^2}$ is the average potential energy of the system. Here $I$ is the rotational inertia of the system and $\kappa$ the torsion constant of the fiber. From his observations Kappler could calculate Boltzmann's constant $k$ and from the relation $N_0 = R/k$ he could obtain Avogadro's number. His values were $k = 1.36 \times 10^{-23}$ joule/molecule K° ± 3% (the accepted value today of $1.380 \times 10^{-23}$ joule/molecule K° being within the limits of error) and $N_0 = 6.1 \times 10^{23}$ particles/mole.

## 24–5 The van der Waals Equation of State

In the preceding chapter we discussed the behavior of an ideal gas. On the macroscopic scale its fundamental relationship is the equation of state

$$pV = \mu RT.$$

From this equation and the principles of thermodynamics we can show that the internal energy $U$ of a gas depends only on the temperature. Real gases obey these relations fairly well at low densities, but their behavior may become markedly different as the density increases. We cannot neglect these deviations from ideal behavior in accurate scientific work. For example, to establish the Kelvin thermodynamic scale in the laboratory we must know how to make the necessary corrections to the scale of a constant-volume gas thermometer. We must therefore know the behavior of real gases rather accurately. Even more important, perhaps, is the fact that the behavior of real gases gives us information on the nature of intermolecular forces and the structure of molecules.

Kinetic theory provides the microscopic description of the behavior of an ideal gas. We have already suggested how the assumptions of kinetic theory could become invalid if applied to a real gas. Under some conditions we may not be justified in neglecting the facts that the molecules occupy a fraction of the volume available to the gas and that the range of molecular forces is greater than the size of the molecule. At high densities we cannot ignore these effects.

J. D. van der Waals (1837–1923) deduced a modified equation of state which takes these factors into account in a simple way. Let us imagine the molecules to be hard spheres of diameter $d$. The diameter of such a sphere would correspond to the distance between the centers of molecules at which strong collision forces come into play. During its motion the center of a molecule cannot approach within a distance $d/2$ from a wall or a distance $d$ from the center of another molecule. Hence the actual volume available to a molecule is smaller than the volume of the containing vessel. Just how much smaller depends on how many molecules there are. Let us represent the volume per mole, $V/\mu$, by $v$. Then the "free volume" per mole would be less than this by the "covolume" $b$. Hence we modify the equation of state from the ideal relation $pv = RT$ to

$$p(v - b) = RT$$

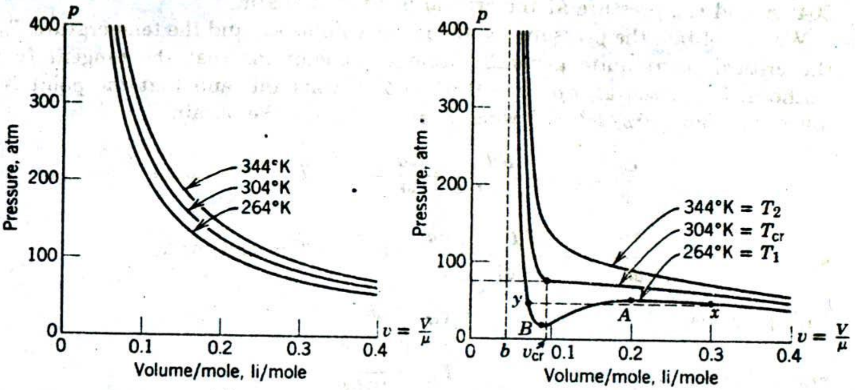
to allow for this. Because of the reduced volume, the number of impacts on the wall increases, thereby increasing the pressure; this relationship was first derived by Clausius.

We can also allow for the effect of attractive forces between molecules in a simple way. Imagine a plane passed through a gas and consider, at any instant, the intermolecular forces which act across it. Each molecule on the left, say, will attract and be attracted by some small number $n$ of those on the right. Now compare this situation with another similar in every way except that the number of molecules per unit volume is doubled. Here any particular molecule on the left will interact on the average with $2n$ of those on the right, for the range of the molecular force is the same, and twice as many molecules now fall into this range. Since there also are twice as many molecules on the left as before which attract in this way, it is clear that the number of attractive pairs across the plane has increased fourfold. Therefore, the effect of these forces varies as the *square* of the number of particles per unit volume or inversely as the square of the volume per mole, that is, as $(1/v)^2$. Because of these intermolecular force bonds, the gas should, for a given external pressure, occupy a volume less than the volume it would occupy as an ideal gas, in which there are no such attractive forces. Or, equivalently, the gas acts as though it is subject to a pressure in excess of the externally applied pressure. This excess pressure is proportional to $(1/v)^2$, or equal to $a/v^2$ where $a$ is a constant. Hence we obtain the *van der Waals equation of state* of a gas,

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT. \tag{24–5}$$

The values of $a$ and $b$ are to be found from experiment, and in this respect the equation is empirical. We must realize that these corrections to the ideal gas equation of state are of the simplest kind, and that failure of the van der Waals equation in any particular case is evidence that our assumptions are oversimplified for that case. No one simple formula is known which applies to all gases under all conditions.

We have seen that real gases do not follow the ideal gas law exactly. Our discussion suggests also that for real gases the internal energy $U$ depends on the volume as well as on the temperature. For if there are (long range) attractive forces between molecules, the potential energy increases as the average distance between molecules increases. Hence, we would expect the internal energy of most real gases to increase slightly with the volume at ordinary temperatures, and this is found to be the case. Of course, collisions can be regarded as arising from repulsive forces. If the molecules move rapidly so as to make many collisions, the potential energy



(a) (b)

**Fig. 24–8** (*a*) Isotherms for an ideal gas. (*b*) Isotherms for a van der Waals gas. We have assumed $a = 3.59$ li²atm/mole² and $b = 0.0427$ li/mole in Eq. 24.4. These values give the best fit of this equation to $p$-$V$-$T$ data for the real gas $CO_2$. $T_{cr}(= 304°\text{ K})$ is the critical temperature.

of the (short range) repulsive forces may be more important than that of the attractive forces and the internal energy could decrease as the volume increases. This is true for hydrogen and helium at ordinary temperatures. In either case, however, the internal energy $U$ is not a function of temperature alone but depends also on the volume. The dependence of the internal energy of a gas on the volume can be deduced readily from the observed results of the free expansion experiment, discussed in Chapter 22.

▶ **Example 4.** On a pressure-volume diagram compare the behavior of an ideal gas at constant temperature to that of a van der Waals gas.

In Fig. 24–8*a* we draw the isotherms (curves of constant $T$) according to the law $pv = RT$. Figure 24–8*b* shows the isotherms according to the law

$$(p + a/v^2)(v - b) = RT.$$

The ideal gas isotherms are each one branch of a rectangular hyperbola, $pv =$ constant. For the van der Waals gas the pressure varies with volume as

$$p = \frac{RT}{(v - b)} - \frac{a}{v^2}. \tag{24–6}$$

As the volume per mole $v$ decreases from large values, the pressure rises, but the $a/v^2$ term, which diminishes the pressure, climbs rapidly so that for sufficiently low $T$ the pressure passes through a maximum at $A$. As $v$ is further decreased, the $RT/(v - b)$ term climbs more rapidly so that the pressure goes through a minimum at $B$ and then rises rapidly without bound as $v$ tends to the value $b$. At neighboring higher temperatures, the maxima and minima are less pronounced and are closer to the inflection point that lies between them. At the so-called critical temperature ($T = T_{cr}$), they coincide in a horizontal inflection point called the critical point. For temperatures sufficiently higher than the critical temperature $T_{cr}$ the van der Waals isotherms have no inflection point and approach the rectangular-hyperbola behavior of the ideal-gas isotherms. For carbon dioxide the critical temperature is 304° K and the pressure at the critical point is 72.9 atm.

We can obtain the pressure $p_{cr}$, the molar volume $v_{cr}$, and the temperature $T_{cr}$ of the critical point quite generally from the conditions that the tangent to the isotherm is horizontal, $dp/dv = 0$ when $T =$ constant, and that the point is an inflection point, $d^2p/dv^2 = 0$ when $T =$ constant. We obtain

$$\frac{dp}{dv} = -\frac{RT}{(v-b)^2} + \frac{2a}{v^3} = 0 \quad (T = \text{constant})$$

and

$$\frac{d^2p}{dv^2} = \frac{2RT}{(v-b)^3} - \frac{6a}{v^4} = 0 \quad (T = \text{constant.})$$

This gives us

$$v_{cr} = 3b$$

and

$$T_{cr} = \frac{8a}{27bR}$$

Putting these in Eq. 24–6, we obtain

$$p_{cr} = \frac{a}{27b^2}.$$

The isotherms suggest the actual experimental behavior of liquids and gases. The maxima and minima of the isotherms below the critical temperature are not usually observed experimentally. At some point $x$ the gas begins to condense. As the volume is decreased, the pressure remains constant (dotted line) until at $y$ all the gas has been transformed into liquid. Beyond $y$, as we decrease the volume, we are compressing a liquid, with the consequent sharp rise in pressure needed to make even small volume changes. Actually the portions $xA$ and $By$ of the isotherms can be obtained experimentally by using very pure gases and liquids. We call these supersaturated vapors and supercooled liquids, and they are in metastable states. The portion $AB$ cannot be reproduced experimentally and is unstable.

The constants $a$ and $b$ in van der Waals equation can be calculated from the experimental values of the critical quantities. The term $a/v^2$ is called an *internal pressure*. Some values for air are of interest. For air at 0° C and external pressure $p$ of 1.00 atm, the internal pressure is 0.0028 atm; at 0°C and external pressure $p$ of 100 atm, the internal pressure is 26 atm. For air at −75° C the corresponding values of the internal pressure are 0.0056 atm and 84.5 atm. When a gas expands under pressure and does work against outside compressing forces, it must also do work against these internal forces. For air at −75° C and 100 atm, the work done against internal forces is nearly as great as that done against external forces. There is an important distinction between internal and external work, however. In the case of external work, energy is transferred from the body to an outside body; in the case of internal work, there is merely a transfer from one kind of energy to another within the body, as from potential to kinetic. The constant $b$ varies from gas to gas, but is usually of the order of 30 cm$^3$/mole. Hence the covolume is about 0.15% of the free volume available to a gas at standard conditions.

Although the van der Waals formula is a good qualitative guide, the quantitative experimental data cannot be matched everywhere with constant values for $a$ and $b$. The reason is that the model on which the formula is based is still an oversimplification. Instead of assuming that the molecules always have a well-defined diameter, for example, we must use the actual intermolecular force (Fig. 23–3). In this way a more accurate correction to the ideal gas law can be made. Van der Waals knew this would be necessary for accurate quantitative work.

## QUESTIONS

**1.** Consider the case in which the mean free path is greater than the longest straight line in a vessel. Is this a perfect vacuum for a molecule in this vessel?

**2.** Give a qualitative explanation of the connection between the mean free path of ammonia molecules in air and the time it takes to smell the ammonia when a bottle is opened across the room.

**3.** The two opposite walls of a container of gas are kept at different temperatures. Describe the mechanism of heat conduction through the gas.

**4.** A gas can transmit only those sound waves whose wavelength is long compared with the mean free path. Can you explain this? Where might this limitation arise?

**5.** If molecules are not spherical, what meaning can we give to $d$ in Eq. 24–1 for the mean free path? In which gases would the molecules act the most nearly as rigid spheres?

**6.** Suppose we dispense with the hypothesis of elastic collisions in kinetic theory and consider the molecules as centers of force acting at a distance. Does the concept of mean free path have any meaning under these circumstances?

**7.** Since the actual force between molecules depends on the distance between them, forces can cause deflections even when molecules are far from "contact" with one another. Furthermore, the deflection caused should depend on how long a time these forces act and hence on the relative speed of the molecules. (*a*) Would you then expect the measured mean free path to depend on temperature, even though the density remains constant? (*b*) If so, would you expect $\bar{l}$ to increase or decrease with temperature? (*c*) How does this dependence enter into Eq. 24–1?

**8.** Justify qualitatively the statement that, in a mixture of molecules of different kinds in complete equilibrium, each kind of molecule has the same Maxwellian distribution in speed that it would have if the other kinds were not present.

**9.** The Maxwellian distribution of speeds among molecules in a gas is shown in Fig. 24–4. How would you expect the Maxwellian distribution of *velocities* to look? What would the average velocity be?

**10.** The fraction of molecules within a given range $\Delta v$ of the root-mean-square speed decreases as the temperature of a gas rises. Explain why.

**11.** (*a*) Do half the molecules in a gas in thermal equilibrium have speeds greater than $v_p$? Than $\bar{v}$? Than $v_{\text{rms}}$?

(*b*) Which speed, $v_p$, $\bar{v}$, or $v_{\text{rms}}$, corresponds to a molecule having average kinetic energy?

**12.** The slit system in Fig. 24–5 selects only those molecules moving in the $+x$-direction. Does this destroy the validity of the experiment as a measure of the distribution of speeds of molecules moving in all directions?

**13.** Why did Rainwater and Havens, in their investigation of the speed distribution of neutrons (page 607), select paraffin as a material to bring fast neutrons rather quickly into thermal equilibrium?

**14.** List examples of the Brownian motion in physical phenomena.

**15.** We have defined $n$ to be the number of molecules per unit volume in a gas. If we define $n$ for a very small volume in a gas, say one equal to ten times the volume of an atom, then $n$ fluctuates with time through the range of values zero to some maximum value. How then can we justify a statement that $n$ has a definite value at every point in the gas?

**16.** Show that as the volume per mole of a gas increases, the van der Waals equation tends to the equation of state of an ideal gas.

**17.** The covolume $b$ in van der Waals equation is often taken to be four times the actual volume of the gas molecules themselves. What factors would have to be taken into account to obtain such a result?

**18.** Keeping in mind that internal energy of a body consists of kinetic energy and potential energy of its particles how would you distinguish between the internal energy of a body and its temperature?

## PROBLEMS

**1.** The mean free path of nitrogen molecules at 0° C and 1 atm is $0.80 \times 10^{-5}$ cm. At this temperature and pressure there are $2.7 \times 10^{19}$ molecules/cm$^3$. What is the molecular diameter?

**2.** The best vacuum attained so far in the laboratory is $10^{-10}$ mm-Hg. How many molecules of gas remain per cubic centimeter at 20° C in this "vacuum"?

**3.** In the cosmotron at the Brookhaven National Laboratory the protons travel around a circular path of diameter 75 ft in a chamber of $10^{-6}$ mm-Hg pressure.

(*a*) Estimate the number of gas molecules per cubic centimeter at this pressure.

(*b*) What is the mean free path of the gas molecules under these conditions if the molecular diameter is $2.0 \times 10^{-8}$ cm?

**4.** At what frequency would the wavelength of sound be of the order of the mean free path in oxygen at 1-atm pressure and 0° C? Take the diameter of the oxygen molecule to be $3.00 \times 10^{-8}$ cm.

**5.** For a gas in which all molecules travel with the same speed $v$, show that $\bar{v}_{\text{rel}} = \frac{4}{3}v$ rather than $\sqrt{2}\,\bar{v}$ (which is the result obtained when we consider the actual distribution of molecular speeds). See p. 601.

**6.** At 2500 km above the earth's surface the density is about one molecule/cm$^3$. What mean free path is predicted by Eq. 24-1 and what is its significance under these conditions?

**7.** The mean free path of a molecule is $\bar{l}$. Prove that the probability that a molecule will go *at least* a distance $x$ before having its next collision is $e^{-x/\bar{l}}$.

**8.** The mean free path $\bar{l}$ of the molecules of a gas may be determined from measurements (e.g., from measurement of the viscosity of the gas). At 20° C and 75 cm-Hg pressure such measurements yield values of $\bar{l}_A$ (argon) $= 9.9 \times 10^{-6}$ cm and $\bar{l}_{N_2}$ (nitrogen) $= 27.5 \times 10^{-6}$ cm. (*a*) Find the ratio of the effective cross-section diameters of argon and nitrogen. (*b*) What would the value be of the mean free path of argon at 20° C and 15 cm-Hg? (*c*) What would the value be of the mean free path of argon at −40° C and 75 cm-Hg?

**9.** A molecule of hydrogen (diameter $10^{-8}$ cm) escapes from a furnace ($T = 4000°$ K) with the root-mean-square speed into a chamber containing atoms of cold argon (diameter $3 \times 10^{-8}$ cm) at a density of $4 \times 10^{19}$ atoms/cm$^3$. (*a*) What is the speed of the hydrogen molecule? (*b*) On a collision between the molecule and an argon atom, what is the closest distance between their centers, considering each as spherical? (*c*) What is the initial number of collisions per unit time experienced by the hydrogen molecule?

**10.** You are given the following group of particles ($N_i$ represents the number of particles which have a speed $v_i$).

| $N_i$ | $v_i$(cm/sec) |
|---|---|
| 2 | 1.00 |
| 4 | 2.00 |
| 6 | 3.00 |
| 8 | 4.00 |
| 2 | 5.00 |

(*a*) Compute the average speed $\bar{v}$. (*b*) Compute the *root-mean-square speed* $v_{\mathrm{rms}}$ (*c*) Among the five speeds shown, which is the *most probable speed* $v_p$ for the entire group?

**11.** Consider the distribution of speeds shown in Fig. 24–9. (*a*) List $v_{\mathrm{rms}}$, $\bar{v}$, and $v_p$ in the order of increasing speed. (*b*) How does this compare with the Maxwellian distribution?

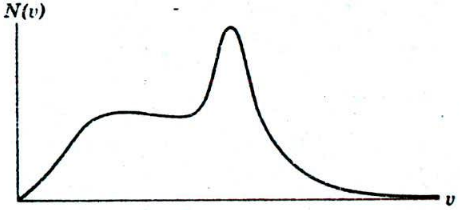


Fig. 24–9

**12.** In the apparatus of Miller and Kusch (Fig. 24–5) the length $l$ of the rotating cylinder is 20.4 cm and the angle $\phi$ is $(2\pi/74.7)$ radians. What rotational speed corresponds to a selected speed $v$ of 200 meters/sec?

**13.** Calculate the root-mean-square speed of smoke particles of mass $5.0 \times 10^{-14}$ gm in air at 0° C and 1-atm pressure.

**14.** Particles of mass $6.2 \times 10^{-14}$ gm are suspended in a liquid at 27° C and are observed to have a root-mean-square speed of 1.4 cm/sec. Calculate Avogadro's number from the equipartition theorem and these data.

**15.** Colloidal particles in solution are buoyed up by the liquid in which they are suspended. Let $\rho'$ be the density of liquid and $\rho$ the density of the particles. If $V$ is the volume of a particle, show that the number of particles per unit volume in the liquid varies with height as

$$n = n_0 \exp\left[-\frac{N_0}{RT} V(\rho - \rho')gh\right].$$

This equation was tested by Perrin in his Brownian motion studies.

**16.** The average speed of hydrogen molecules at 0° C is 1694 meters/sec. Compute the average speed of colloidal particles of "molecular weight" $3.2 \times 10^6$ gm/mole.

**17.** Calculate the work done per mole in an isothermal expansion of a van der Waals gas from volume $V_i$ to $V_f$.

**18.** The constant $a$ in van der Waals equation is 0.37 nt-m$^4$/mole$^2$ for $CO_2$ and 0.025 nt-m$^4$/mole$^2$ for hydrogen. Compute the internal pressures for these gases for values of $v/v_0$ (where $v_0 = 22.4$ liters/mole) of 1, 0.01, and 0.001.

**19.** (*a*) The constant $b$ in van der Waals equation is 43 cm$^3$/mole for $CO_2$. Using the value for $a$ in the previous problem, compute the pressure at 0° C for a specific volume of 0.55 liter/mole, assuming van der Waals equation to be strictly true. (*b*) What is the pressure under these same conditions, assuming $CO_2$ behaves as an ideal gas?

**20.** Van der Waals' $b$ for oxygen is 32 cm$^3$/mole. Assume $b$ is four times the actual volume of a mole of "billiard-ball" $O_2$ molecules and compute the diameter of an $O_2$ molecule.

**21.** The constants $a$ and $b$ in the van der Waals equation are different for different substances. Show, however, that if we take $v_{cr}$, $p_{cr}$, and $T_{cr}$ as the units of specific volume, pressure, and temperature, the van der Waals equation becomes identical for all substances.

# Entropy and the Second Law of Thermodynamics

## CHAPTER 25

### 25–1 Introduction

The first law of thermodynamics states that energy is conserved. However, we can think of many thermodynamic processes which conserve energy but which actually never occur. For example, when a hot body and a cold body are put into contact, it simply does not happen that the hot body gets hotter and the cold body colder. Or again, a pond does not suddenly freeze on a hot summer day by giving up heat to its environment. *And yet neither of these processes violates the first law of thermodynamics.* Similarly, the first law does not restrict our ability to convert work into heat or heat into work, except that energy must be conserved in the process. And yet in practice, although we can convert a given quantity of work completely into heat, we have never been able to find a scheme that converts a given amount of heat completely into work. The second law of thermodynamics deals with this question of whether processes, assumed to be consistent with the first law, do or do not occur in nature. Although the ideas contained in the second law may seem subtle or abstract, in application they prove to be extremely practical.

### 25–2 Reversible and Irreversible Processes

Consider a typical system in thermodynamic equilibrium, say a mass $\mathcal{M}$ of a (real) gas confined in a cylinder-piston arrangement of volume $V$, the gas having a pressure $p$ and a temperature $T$. In an equilibrium state these thermodynamic variables remain constant with time. Suppose that the cylinder, whose walls are an (ideal) heat insulator but whose base is an (ideal) heat conductor is placed on a large heat reservoir maintained at this same temperature $T$, as in Fig. 22-9. Now let us change the system to another equilibrium state in which the temperature $T$ is the same but the volume $V$ is reduced by one-half. Of the many ways in which we could do this we discuss two extreme cases.

I. We depress the piston very rapidly; we then wait for equilibrium with the reservoir to be re-established. During this process the gas is turbulent and its pressure and temperature are not well defined; we cannot plot the process as a continuous line on a $p$-$V$ diagram because we would not know what value of pressure (or temperature) to associate with a given volume. The system passes from one equilibrium state $i$ to another $f$ through a series of nonequilibrium states (Fig. 25-1*a*).

II. We depress the piston (assumed to be frictionless) exceedingly slowly—perhaps by adding sand to the top of the piston—so that the pressure, volume, and temperature of the gas are, at all times, well-defined quantities. We first drop a few grains of sand on the piston. This will reduce the volume of the system a little and the temperature will tend to rise; the system will depart from equilibrium, but only slightly. A small amount of heat will be transferred to the reservoir and in a short time the system will reach a new equilibrium state, its temperature again being that of the reservoir. Then we drop a few more grains of sand on the piston, reducing the volume further. Again we wait for a new equilibrium state to be established, and so forth. By many repetitions of this procedure we finally reduce the volume by one-half. During this entire process the system is never in a state differing much from an equilibrium state. If we imagine carrying out this procedure with still smaller successive increases in pressure, the intermediate states will depart from equilibrium even less. By indefinitely increasing the number of changes and corre-

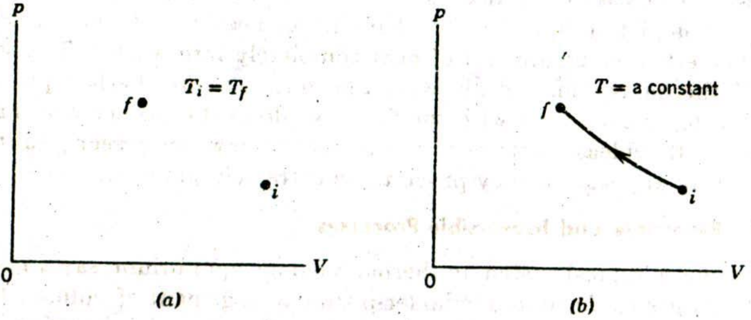


**Fig. 25–1** We cause a real gas to go from an initial equilibrium state $i$ described by $p_i$, $V_i$, $T_i$ to a final equilibrium state $f$ described by $p_f$, $V_f$ $(= \frac{1}{2}V_i)$, and $T_f$ $(= T_i)$. We carry out the process (*a*) irreversibly, and (*b*) reversibly.

spondingly decreasing the size of each change, we arrive at an ideal process in which the system passes through a continuous succession of equilibrium states, which we can plot as a continuous line on a $p$-$V$ diagram (Fig. 25–1$b$). During this process a certain amount of heat $Q$ is transferred from the system to the reservoir.

Processes of type I are called *irreversible* and those of type II are called *reversible*. *A reversible process is one that, by a differential change in the environment, can be made to retrace its path.* Thus as we cause the piston to move slowly downward, in II, the external pressure on the piston exceeds the pressure exerted on it by the gas by only a differential amount $dp$. If at any instant we reduce the external pressure ever so slightly (by removing a few sand grains), so that it is *less than* the internal gas pressure by $dp$, the gas will expand instead of contracting and the system will retrace the equilibrium states through which it has just passed.* In practice all processes are irreversible, but we can approach reversibility arbitrarily closely by making appropriate experimental refinements. The strictly reversible process is a simple and useful abstraction that bears a similar relation to real processes that the ideal gas abstraction does to real gases.

The process described in II is not only reversible but *isothermal*, because we have assumed that the temperature of the gas differs at all times by only a differential amount $dT$ from the (constant) temperature of the reservoir on which the cylinder rests.

We could also reduce the volume *adiabatically* by removing the cylinder from the heat reservoir and putting it on a nonconducting stand. In an adiabatic process no heat is allowed to enter or to leave the system. An adiabatic process can be either reversible or irreversible—the definition does not exclude either. In a reversible adiabatic process we move the piston exceedingly slowly—perhaps using the sand-loading technique; in an irreversible adiabatic process we shove the piston down quickly.

The temperature of the gas will rise during an adiabatic compression because, from the first law, with $Q = 0$, the work $W$ done in pushing down the piston must appear as an increase $\Delta U$ in the internal energy of the system. $W$ will have different values for different rates of pushing down the piston, being given by $\int p\,dV$—that is, by the area under a curve on a $p$-$V$ diagram—only for reversible processes (for which $p$ has a well-defined value). Thus $\Delta U$ and the corresponding temperature change $\Delta T$ will not be the same for reversible and irreversible adiabatic processes.

---

* Not all processes carried out very slowly are reversible. For example, if the piston in our example exerted a frictional force against the cylinder walls, it would not reverse its motion if we made only a differential change $dp$ in the external pressure. We would have to make a change $\Delta p$ that might be an appreciable fraction of $p$. Thus our criterion for reversibility, which involves a response of the system to a *differential* change in the environment, is not met. The word *quasi-static* is used to describe processes that are carried out slowly enough so that the system passes through a continuous sequence of equilibrium states; a quasi-static process may or may not be reversible. See "Thermodynamics of an Irreversible Quasi-Static Process," by John S. Thomsen, *American Journal of Physics*, **28**, 119, 1960.

## 25–3 The Carnot Cycle

Suppose that we have a system (a real gas, say) in an equilibrium state in a cylinder-piston arrangement. By using our ability to make changes in the environment of the system we can carry out, at our pleasure, a wide variety of processes. We can let the gas expand or we can compress it; we can add or subtract energy in the form of heat; we can do these things and others irreversibly or reversibly. We can also choose to carry out a sequence of processes such that the system returns to its original equilibrium state; we call this a *cycle*. If the processes involved are all reversible, we call it a *reversible cycle*.

Figure 25–2 shows a reversible cycle on a $p$-$V$ diagram. Along the curve *abc* we allow the system to expand, and the area under this curve represents the work done by the system during the expansion. Along the curve *cda*, which returns the system to its original state, we compress the system, and the area under this curve represents the work we must do on the system during the compression. Hence, the net work done by the system is represented by the area enclosed by the curve and is positive. If we decided to traverse the cycle in the opposite sense, that is, expanding along *adc* and compressing along *cba*, the net work done by the system would be the negative of that of the previous case.

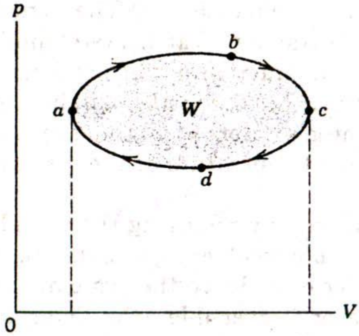


**Fig. 25–2** A $p$-$V$ diagram of a gas undergoing a reversible cycle. The shaded area $W$ represents the net work done by the gas in the cycle.

An important reversible cycle is the *Carnot cycle*, introduced by Sadi Carnot in 1824. We shall see later that this cycle will determine the limit of our ability to convert heat into work. The system consists of a "working substance," such as a gas, and the cycle is made up of two isothermal and two adiabatic reversible processes. The working substance, which we can think of as an ideal gas for concreteness, is contained in a cylinder with a heat-conducting base and nonconducting walls and piston. We also provide, as part of the environment, a heat reservoir in the form of a body of large heat capacity at a temperature $T_1$, another reservoir of large heat capacity at a temperature $T_2$, and two nonconducting stands. We carry out the Carnot cycle in four steps, as shown in Fig. 25–3. The cycle is shown on the $p$-$V$ diagram of Fig. 25–4.

*Step 1.* The gas is in an initial equilibrium state represented by $p_1$, $V_1$, $T_1$ (*a*, Fig. 25–4). We put the cylinder on the heat reservoir at tempera-

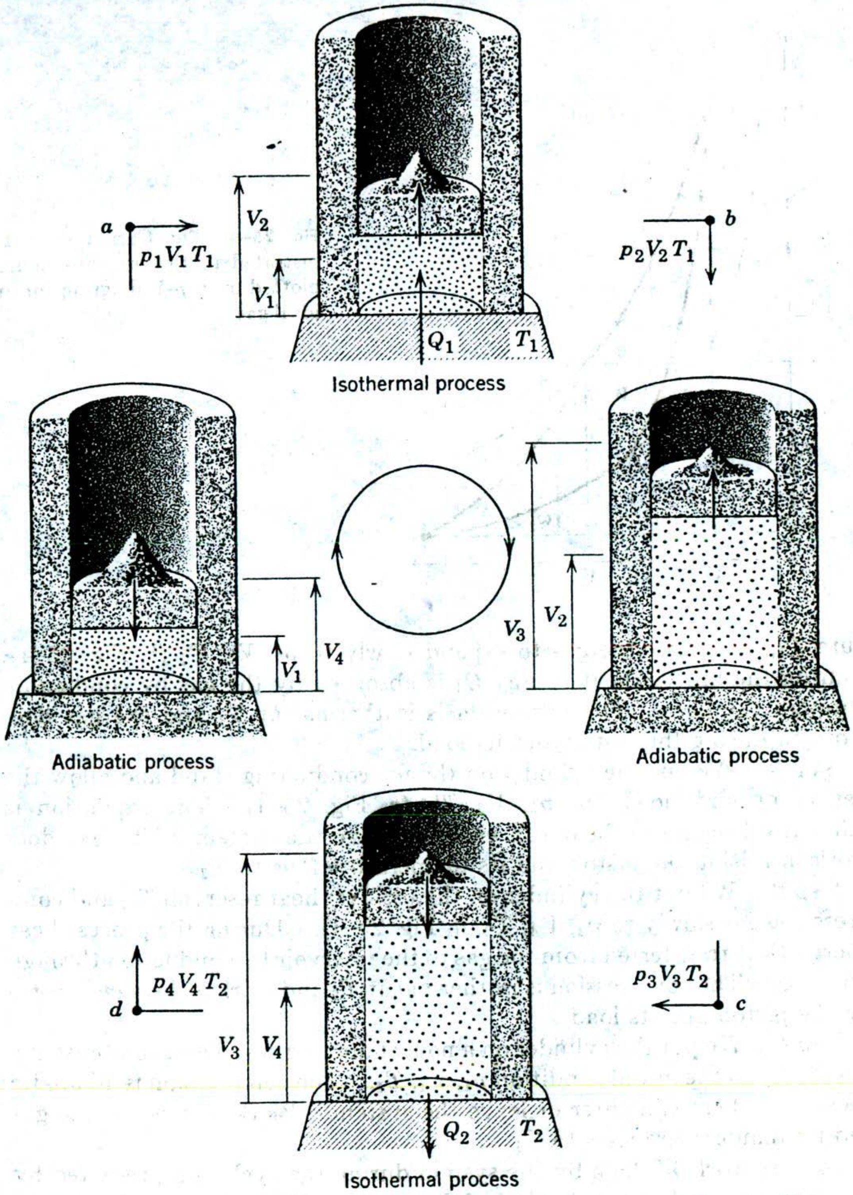


**Fig. 25-3** A Carnot cycle. The points *a*, *b*, *c* and *d* correspond to the points so labelled in Fig. 25.4. The cylinder-piston arrangements show intermediate steps in the processes that connect adjacent points. The arrows on the pistons suggest expansions (caused by removing sand) and compressions (caused by adding sand).

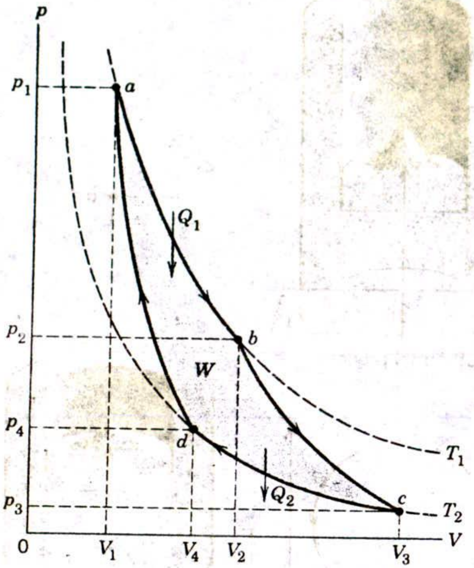


**Fig. 25–4** The Carnot cycle illustrated in the previous figure, plotted on a $p$-$V$ diagram for an ideal gas.

ture $T_1$, and allow the gas to expand slowly to $p_2$, $V_2$, $T_1$ ($b$, Fig. 25–4). During the process heat energy $Q_1$ is absorbed by the gas by conduction through the base. The expansion is isothermal at $T_1$ and the gas does work in raising the piston and its load.

*Step 2.* We put the cylinder on the nonconducting stand and allow the gas to expand slowly to $p_3$, $V_3$, $T_2$ ($c$, Fig. 25–4). The expansion is adiabatic because no heat can enter or leave the system. The gas does work in raising the piston and its temperature falls to $T_2$.

*Step 3.* We put the cylinder on the (colder) heat reservoir $T_2$ and compress the gas slowly to $p_4$, $V_4$, $T_2$ ($d$, Fig. 25–4). During the process heat energy $Q_2$ is transferred from the gas to the reservoir by conduction through the base. The compression is isothermal at $T_2$ and work is done *on* the gas by the piston and its load.

*Step 4.* We put the cylinder on a nonconducting stand and compress the gas slowly to the initial condition $p_1$, $V_1$, $T_1$. The compression is adiabatic because no heat can enter or leave the system. Work is done on the gas and its temperature rises to $T_1$.

The net work $W$ done by the system during the cycle is represented by the area enclosed by path $abcd$ of Fig. 25–4. The net amount of heat energy received by the system in the cycle is $Q_1 - Q_2$, where $Q_1$ is the heat absorbed in Step 1 and $Q_2$ is that given up in Step 3. The initial and final states are the same so that there is no net change in the internal energy $U$ of the system. Hence, from the first law of thermodynamics,

$$W = Q_1 - Q_2 \tag{25–1}$$

for the cycle, in which $Q_1$ and $Q_2$ are taken as positive quantities. The result of the cycle is that heat has been converted into work by the system. Any required amount of work can be obtained by simply repeating the cycle. Hence, the system acts like a *heat engine.*

We have used an ideal gas as an example of a working substance. The working substance can be anything at all, although the $p$-$V$ diagrams for other substances would be different. Common heat engines use steam or a mixture of fuel and air, or fuel and oxygen as their working substance. Heat may be obtained from the combustion of a fuel such as gasoline or coal, or from the annihilation of mass in nuclear fission processes in nuclear reactors. Heat may be discharged at the exhaust or to a condenser. Although real heat engines do not operate on a reversible cycle, the Carnot cycle, which is reversible, gives useful information about the behavior of any heat engine.

The efficiency $e$ of a heat engine is the ratio of the net work done by the engine during one cycle to the heat taken in from the high temperature source in one cycle.* Hence,

$$e = \frac{W}{Q_1} = \frac{Q_1 - Q_2}{Q_1} = 1 - \frac{Q_2}{Q_1}. \tag{25-2}$$

Equation 25-2 shows that the efficiency of a heat engine is less than one (100%) so long as the heat $Q_2$ delivered to the exhaust is not zero. Experience shows that every heat engine rejects some heat during the exhaust stroke. This represents the heat absorbed by the engine that is not converted to work in the process.

We may choose to carry out the Carnot cycle by starting at any point, such as $a$ in Fig. 25-4, and traversing each process in a direction opposite to that of the arrowheads in that figure. Then an amount of heat $Q_2$ is *removed* from the lower temperature reservoir at $T_2$, and an amount of heat $Q_1$ is *delivered* to the higher temperature reservoir at $T_1$; work must be done *on* the system by an outside agency. In this reversed cycle work must be done *on* the system which extracts heat from the lower temperature reservoir. Any amount of heat can be removed from this reservoir by simply repeating the reverse cycle. Hence, the system acts like a *refrigerator*, transferring heat from a body at a lower temperature (the freezing compartment) to one at a higher temperature (the room) by means of work supplied to it (the electric power input).

▶ **Example 1.** Show that the efficiency of a Carnot engine using an ideal gas as the working substance is $e = (T_1 - T_2)/T_1$.

Along the isothermal path $ab$, the temperature, and hence the internal energy of an ideal gas, remains constant. From the first law, the heat $Q_1$ absorbed by the gas in its expansion must be equal to the work $W_1$ done in this expansion. From

---

* The definition reflects the economic importance of engines. Work $W$ is the desirable output; the heat $Q_1$, is the input that must be paid for in the form, say, of a fuel bill. An efficient engine has a large ratio of $W$ to $Q_1$.

Example 2, Chapter 23, we have,

$$Q_1 = W_1 = \mu R T_1 \ln (V_2/V_1).$$

Likewise, in the isothermal compression along the path *cd*, we have

$$Q_2 = W_2 = \mu R T_2 \ln (V_3/V_4).$$

On dividing the first equation by the second, we obtain

$$\frac{Q_1}{Q_2} = \frac{T_1 \ln (V_2/V_1)}{T_2 \ln (V_3/V_4)}.$$

From the equation describing an isothermal process for an ideal gas we obtain for the paths *ab* and *cd*

$$p_1 V_1 = p_2 V_2,$$

$$p_3 V_3 = p_4 V_4.$$

From the equation describing an adiabatic process for an ideal gas we have for paths *bc* and *da*

$$p_2 V_2^\gamma = p_3 V_3^\gamma,$$

$$p_4 V_4^\gamma = p_1 V_1^\gamma.$$

Multiplying these four equations together and canceling the factor $p_1 p_2 p_3 p_4$ appearing on both sides, we obtain

$$V_1 V_2^\gamma V_3 V_4^\gamma = V_2 V_3^\gamma V_4 V_1^\gamma,$$

from which

$$(V_2 V_4)^{\gamma - 1} = (V_3 V_1)^{\gamma - 1}$$

and

$$V_2/V_1 = V_3/V_4.$$

Using this result in our expression for $Q_1/Q_2$, we see that

$$Q_1/Q_2 = T_1/T_2, \tag{25-3}$$

so that

$$e = 1 - Q_2/Q_1 = 1 - T_2/T_1$$

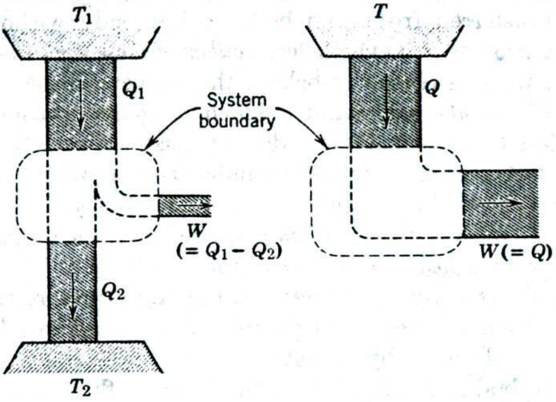
or

$$e = \frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1}.$$

The temperatures $T_1$ and $T_2$ are those measured on the ideal gas scale described in Chapter 21. ◀

## 25–4 The Second Law of Thermodynamics

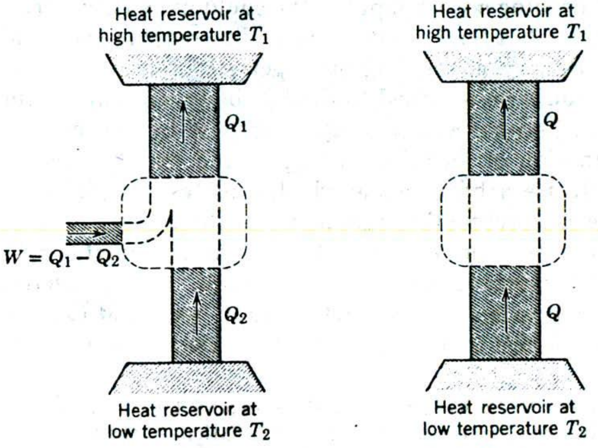
The first heat engines constructed were very inefficient devices. Only a small fraction of the heat absorbed at the high-temperature source could be converted to useful work. Even as engineering design improved, a sizable fraction of the absorbed heat was still discharged at the lower-temperature exhaust of the engine, remaining unconverted to mechanical energy. It remained a hope to devise an engine that could take heat from an abundant reservoir, like the ocean, and convert it completely into useful work. Then it would not be necessary to provide a source of heat at a higher temperature than the outside environment by burning fuels (Fig.



Actual heat engine "Perfect" heat engine

**Fig. 25-5** In an actual heat engine, some of the heat $Q_1$ taken in by the engine is converted into work $W$, but the rest is rejected as heat $Q_2$. In a "perfect" heat engine all the heat input would be converted into work output.

25-5). Likewise, we might hope to be able to devise a refrigerator that simply transfers heat from a cold body to a hot body, without requiring the expense of outside work (Fig. 25-6). *Neither of these hopeful ambitions violates the first law of thermodynamics.* The heat engine would simply convert heat energy completely into mechanical energy, the total energy being



Actual refrigerator "Perfect" refrigerator

**Fig. 25-6** In an actual refrigerator, work $W$ is needed to transfer heat from a low-temperature to a high-temperature reservoir. In a "perfect" refrigerator, heat would flow from the low-temperature to the high-temperature reservoir without any work being done on the engine.

conserved in the process. In the refrigerator, the heat energy would simply be transferred from cold body to hot body without any loss of energy in the process. Nevertheless *neither of these ambitions has ever been achieved*, and there is reason to believe they never will be.

The *second law of thermodynamics*, which is a generalization of experience, is an assertion that such devices do not exist. There have been many statements of the second law, each emphasizing another facet of the law, but all can be shown to be equivalent to one another. Clausius stated it as follows: *It is impossible for any cyclical machine to produce no other effect than to convey heat continuously from one body to another at a higher temperature.* This statement rules out our ambitious refrigerator, for it implies that to convey heat continuously from a cold to a hot object it is necessary to supply work by an outside agent. We know from experience that when two bodies are in contact, heat energy flows from the hot body to the cold body. The second law rules out the possibility of heat energy flowing from cold to hot body in such a case and so determines the direction of transfer of heat. The direction can be reversed only by an expenditure of work.

Kelvin (with Planck) stated the second law in words equivalent to these: *A transformation whose only final result is to transform into work heat extracted from a source which is at the same temperature throughout is impossible.** This statement rules out our ambitious heat engine, for it implies that we cannot produce mechanical work by extracting heat from a single reservoir without returning any heat to a reservoir at a lower temperature.

To show that the two statements are equivalent we need to show that, if either statement is false, the other statement must be false also. Suppose Clausius' statement were false so that we could have a refrigerator operating without needing a work input. We could use an ordinary engine to remove heat from a hot body, to do work and to return part of the heat to a cold body. But by connecting our "perfect" refrigerator into the system, this heat would be returned to the hot body without expenditure of work and would become available again for use by the heat engine. Hence, the combination of an ordinary engine and the "perfect" refrigerator would constitute a heat engine which violates the Kelvin-Planck statement. Or we can reverse the argument. If the Kelvin-Planck statement were incorrect, we could have a heat engine which simply takes heat from a source and converts it completely into work. By connecting this "perfect" heat engine to an ordinary refrigerator, we could extract heat from the hot body, convert it completely to work, use this work to run the

---

* This statement needs to be supplemented if we extend thermodynamics to the region of negative Kelvin temperatures. All other formulations of the second law, and indeed, all other laws of thermodynamics apply to negative temperatures without revision. See an article, "Thermodynamics and Statistical Mechanics at Negative Absolute Temperatures," by N. F. Ramsey, in *Temperature, Its Measurement and Control in Science and Industry*, Vol. 3, Part 1, Reinhold Publishing Co., New York, 1962.

ordinary refrigerator, extract heat from the cold body, and deliver it plus the work converted to heat by the refrigerator to the hot body. The net result is a transfer of heat from cold to hot body without expenditure of work and this violates Clausius' statement.

The second law tells us that many processes are irreversible. For example, Clausius' statement specifically rules out a simple reversal of the process of heat transfer from hot body to cold body. Not only will some processes not run backward by themselves, but no combination of processes can undo the effect of an irreversible process without causing another corresponding change elsewhere. In later sections we shall develop these ideas more fully and formulate the second law quantitatively.

## 25–5 The Efficiency of Engines

Carnot first wrote scientifically on the theory of heat engines. In 1824 he published *Reflections on the Motive Power of Heat*. By then the steam engine was commonly used in industry. Carnot wrote:

> In spite of labor of all sorts expended on the steam engine, and in spite of the perfection to which it has been brought, its theory is very little advanced. . . .
>
> The production of motion in the steam engine is always accompanied by a circumstance which we should particularly notice. This circumstance is the passage of caloric from one body where the temperature is more or less elevated to another where it is lower. . . .
>
> The motive power of heat is independent of the agents employed to develop it; its quantity is determined solely by the temperature of the bodies between which, in the final result, the transfer of the caloric occurs.

Hence, Carnot directed attention to the facts that the difference in temperature was the real source of "motive power," that the transfer of heat played a significant role, and that the choice of working substance was of no theoretical importance.

Carnot's achievement was remarkable when we recall that the mechanical equivalence of heat and the conservation of energy principle were not known in 1824. In his later papers, published posthumously in 1872, it became clear that Carnot had foreseen the principle of the conservation of energy and had made an accurate determination of the mechanical equivalent of heat. He had planned a program of research which included all the important developments in the field made by other investigators during the following several decades. However, he died during a cholera epidemic in 1832 at the age of 36, leaving it to others to extend his work. It was William Thomson (later Lord Kelvin) who modified Carnot's reasoning to bring it into accord with the mechanical theory of heat, and who, together with Clausius, successfully developed the science of thermodynamics.

Carnot developed the concept of a reversible engine and the reversible cycle named after him. He stated a theorem of great practical importance: *The efficiency of all reversible engines operating between the same two temperatures is the same and no irreversible engine working between the same two tem-*

*peratures can have a greater efficiency than this.* Clausius and Kelvin showed that this theorem was a necessary consequence of the second law of thermodynamics. Notice that nothing is said about the working substance, so that the efficiency of a reversible engine is independent of the working substance and depends only on the temperatures. Furthermore, a reversible engine operates at the maximum efficiency possible for any engine working between the same two temperature limits. The proof of this theorem follows.

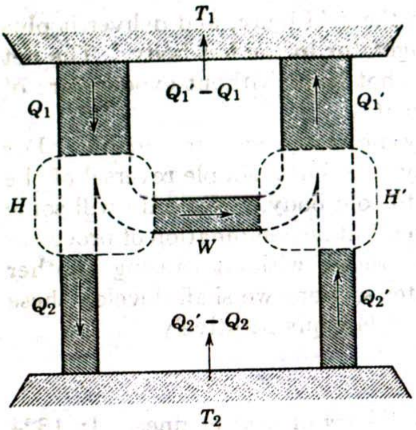


**Fig. 25–7** Proof of Carnot's theorem.

Let us call the two reversible engines $H$ and $H'$. They operate between the temperatures $T_1$ and $T_2$ where $T_1 > T_2$. They may differ, say, in their working substance or in their initial pressures and lengths of stroke. We choose $H$ to run forward and $H'$ to run backward (as a refrigerator). The forward-running engine $H$ takes in heat energy $Q_1$ at $T_1$ and gives out heat energy $Q_2$ at $T_2$. The backward-running engine (refrigerator) $H'$ takes in heat $Q_2'$ at $T_2$ and gives out heat $Q_1'$ at $T_1$. We now connect the engines mechanically and adjust the stroke lengths so that the work done per cycle by $H$ is just sufficient to operate $H'$ (Fig. 25–7). Suppose the efficiency $e$ of $H$ were greater than the efficiency $e'$ of $H'$. Then

$$e > e', \qquad \text{(assumption)}$$

or

$$\frac{Q_1 - Q_2}{Q_1} > \frac{Q_1' - Q_2'}{Q_1'}.$$

Since the work per cycle done by one engine equals the work per cycle done on the other engine,

$$W = W',$$

or

$$Q_1 - Q_2 = Q_1' - Q_2'.$$

Comparing these relations, we see that (since $Q_1 - Q_2 > 0$)

$$\frac{1}{Q_1} > \frac{1}{Q_1'}$$

or

$$Q_1 < Q_1'.$$

Hence (from the work equality),

$$Q_2 < Q_2'.$$

Thus, the hot source gains heat $Q_1' - Q_1$ (positive) and the cool source loses heat $Q_2' - Q_2$ (positive). But no work is done in the process by the combined system $H + H'$ so that we have transferred heat from a body at one temperature to a body at a higher temperature without performing work—in direct contradiction to Clausius' statement of the second law. Hence, we conclude that $e$ cannot be greater than $e'$. Likewise, by reversing the engines we can use the same reasoning to prove that $e'$ cannot be greater than $e$. Hence,

$$e = e',$$

proving the first part of Carnot's theorem.

Now suppose that $H$ is an *irreversible* engine. Then by the exact same procedure we can prove that $e_{ir}$ cannot be greater than $e'$. But $H$ cannot be reversed, so we cannot prove that $e'$ cannot be greater than $e_{ir}$. Therefore, $e_{ir}$ is either equal to or less than $e'$. Since $e' = e = e_{\text{reversible}}$, we have

$$e_{\text{irreversible}} \leqslant e_{\text{reversible}},$$

thus proving the second part of Carnot's theorem.

▶ **Example 2.** A steam engine takes steam from the boiler at 200° C (225 lb/in.$^2$ pressure) and exhausts directly into the air (14 lb/in.$^2$ pressure) at 100° C. What is its maximum possible efficiency?

Using the result of Example 1 (which applies to this case by virtue of Carnot's theorem, which we have just proved) we have

$$e = \frac{T_1 - T_2}{T_1} = \frac{473^\circ\text{ K} - 373^\circ\text{ K}}{473^\circ\text{ K}} \times 100\% = 21.1\%.$$

Actual efficiencies of about 15% are usually realized. Energy is lost by friction, turbulence, and heat conduction. Lower exhaust temperatures on more complicated steam engines may raise the maximum possible efficiency to 35% and the actual efficiency to 20%. The efficiency of an ordinary automobile engine is about 22% and that of a large Diesel oil engine about 40%. ◀

## 25-6 The Thermodynamic Temperature Scale

The efficiency of a reversible engine is independent of the working substance and depends only on the two temperatures between which the engine works. Since $e = 1 - Q_2/Q_1$, then $Q_2/Q_1$ can depend only on the temperatures. This led Kelvin to suggest a new scale of temperature. If we let $\theta_1$ and $\theta_2$ represent these two temperatures, his defining equation is

$$\theta_1/\theta_2 = Q_1/Q_2.$$

That is, two temperatures on this scale are to each other as the heats absorbed and rejected, respectively, by a Carnot engine operating between these temperatures. Such a temperature scale is called the *thermodynamic* (or *Kelvin*) *temperature scale.*

To complete the definition of the thermodynamic scale, we assign the arbitrary value of 273.16° to the temperature of the triple point of water. Hence, $\theta_{tr} = 273.16^\circ$ K. Then for a Carnot engine operating between reservoirs at the temperatures $\theta$ and $\theta_{tr}$ we have

$$\frac{\theta}{\theta_{tr}} = \frac{Q}{Q_{tr}}$$

or

$$\theta = 273.16^\circ\text{ K}\,\frac{Q}{Q_{tr}} \tag{25-4}$$

If we compare this with the corresponding equation for the ideal gas temperature $T$, namely

$$T = 273.16^\circ\text{ K} \lim_{p_{tr}\to 0} \frac{p}{p_{tr}} \tag{25-5}$$

we see that on the thermodynamic scale $Q$ plays the role of a thermometric property. However, $Q$ does not depend on the characteristics of any substance because a Carnot engine is independent of the nature of the working substance. Therefore, we obtain a scale of temperature which is free of the objection we can raise to the ideal gas scale of Chapter 21, and in fact we arrive at a fundamental definition of temperature.

The definition of thermodynamic temperature enables us to rewrite the equation for the efficiency of a reversible engine as

$$e = \frac{Q_1 - Q_2}{Q_1} = \frac{\theta_1 - \theta_2}{\theta_1}. \tag{25-6}$$

But we have shown (Example 1) that the efficiency of a Carnot engine using an ideal gas as working substance is

$$e = \frac{Q_1 - Q_2}{Q_1} = \frac{T_1 - T_2}{T_1} \tag{25-7}$$

where $T$ is the temperature given by the constant-volume thermometer containing the ideal gas. Hence, $Q_1/Q_2 = T_1/T_2$ and $Q_1/Q_2 = \theta_1/\theta_2$. Since $\theta_{tr} = T_{tr} = 273.16°$ and $\theta/\theta_{tr} = T/T_{tr}$, it follows that $\theta = T$. Hence, *if an ideal gas were available for use in a constant-volume thermometer, the thermometer would yield the thermodynamic* (or *Kelvin*) *temperature.* We have seen that, although an ideal gas is not available, measurements made using the limiting process of Eq. 25–5 with real gases correspond to ideal gas behavior. We shall treat the ideal gas scale and the thermodynamic scale as identical and we shall use the designation °K interchangeably for each, as in fact we have already done.

In practice, we cannot have a gas below 1° K. One of the methods used in measuring temperature below 1° K employs the thermodynamic scale directly. The ratio of two thermodynamic temperatures is the ratio of two heats transferred during two isothermal processes bounded by the same two adiabatics (Fig. 25–8). The location of the adiabatic boundaries (on the $p$-$V$ diagram) can be found experimentally, and the heats transferred during two nearly reversible isothermal processes can be measured with great precision.

From the equations

$$T = 273.16°\text{ K}\,\frac{Q}{Q_{tr}} \qquad \text{or} \qquad \frac{T}{T_{tr}} = \frac{Q}{Q_{tr}}$$

it is clear that the heat $Q$ transferred in an isothermal process between two given adiabatics decreases as the temperature $T$ decreases. Conversely, the smaller $Q$ is the lower the corresponding temperature $T$ is. Now the smallest possible value of $Q$ is zero and the corresponding $T$ is absolute zero. That is, *if a system undergoes a reversible isothermal process with no transfer of heat, the temperature at which this process takes place is the absolute zero.* Hence, at absolute zero, an isothermal and an adiabatic process are identical (Fig. 25–8).

This definition of absolute zero applies to all substances and is independent of the properties of any one of them. Notice that no reference is made to molecules or molecular energy and that we have obtained a purely macroscopic definition of absolute zero.

The efficiency of a Carnot engine is

$$e = 1 - \frac{T_2}{T_1},$$

which is the maximum possible efficiency any engine can have operating between temperatures $T_1$ and $T_2$. To obtain 100% efficiency, $T_2$ must be zero. Only when the low-temperature reservoir is at absolute zero will all the heat absorbed at the high-temperature reservoir be converted to work.

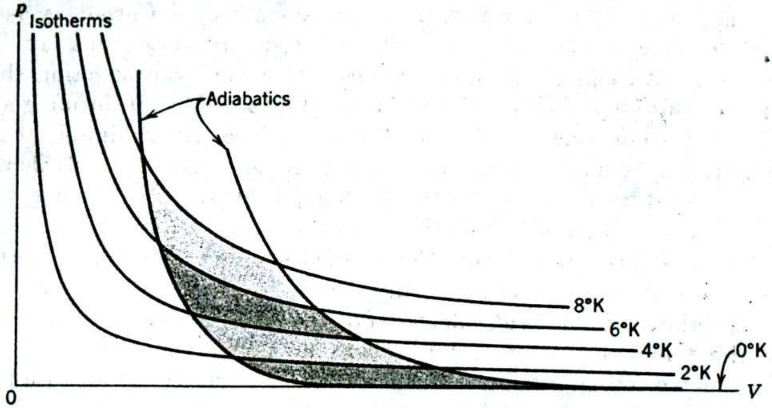


**Fig. 25–8** A series of Carnot cycles tending toward absolute-zero temperature, as used in establishing the thermodynamic scale of temperature. The difference in slope between isothermals and adiabatics has here been exaggerated for clarity.

The fundamental feature of all cooling processes is that the lower the temperature, the more difficult it is to go still lower. This experience has led to the formulation of the *third law of thermodynamics,* which can be stated in one form as follows: *It is impossible by any procedure, no matter how idealized, to reduce any system to the absolute zero of temperature in a finite number of operations.* Hence, because we cannot obtain a reservoir at absolute zero, a heat engine with 100% efficiency is a practical impossibility.

## 25–7 Entropy—Reversible Processes

The zeroth law of thermodynamics is related to the concept of *temperature T* and the first law is related to the concept of *internal energy U.* In this and the following sections we show that the second law of thermodynamics is related to a thermodynamic variable called *entropy, S,* and that we can express the second law quantitatively in terms of this variable. We start by considering a Carnot cycle. For such a cycle we have seen (Eq. 25–3) that

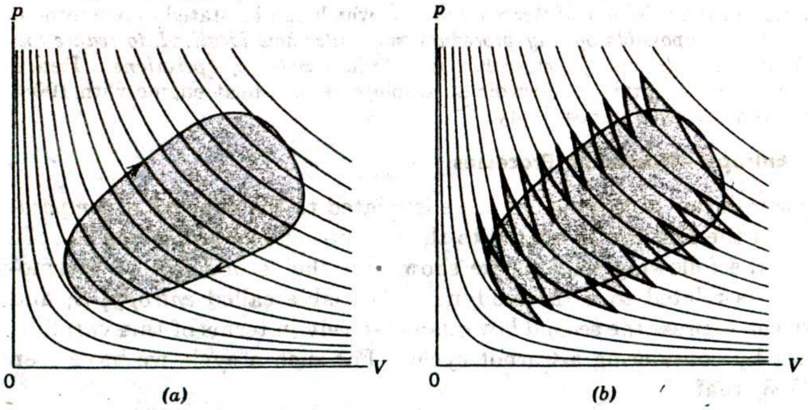
$$\frac{Q_1}{T_1} = \frac{Q_2}{T_2},$$

in which the $Q$'s were taken as positive quantities, that is, we dealt with the magnitudes, or absolute values, only of the $Q$'s. If we now interpret them again as algebraic quantities, $Q$ being positive when heat enters the system and negative when heat leaves the system, we can write this relation as

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0.$$

This equation states that the sum of the algebraic quantities $Q/T$ is zero for a Carnot cycle.

As a next step, we assert that *any* reversible cycle is equivalent, to as close an approximation as we wish, to an assembly of Carnot cycles. Figure 25–9*a* shows an arbitrary reversible cycle superimposed on a family of isotherms. We can approximate the actual cycle by connecting the isotherms by suitably chosen adiabatic lines (Fig. 25–9*b*), thus forming an assembly of Carnot cycles. The student should convince himself that traversing the individual Carnot cycles in Fig. 25–9*b* is exactly equivalent, in terms of heat transferred and work done, to traversing the jagged sequence of isotherms and adiabatic lines that approximates the actual cycle. This is so because adjacent Carnot cycles have a common isotherm and the two traversals, in opposite directions, cancel each other in the region of overlap as far as heat transfer and work done are concerned. By making the temperature interval between the isotherms in Fig. 25–9*b* small enough we can approximate the actual cycle as closely as we wish by an alternating sequence of isotherms and adiabatic lines.



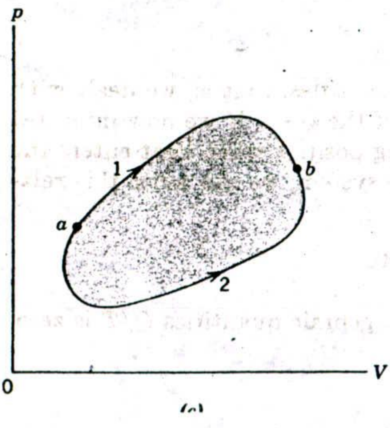


**Fig. 25–9** (*a*) A reversible cycle superimposed on a family of isotherms. (*b*) The isotherms are connected by adiabatic lines, forming an assembly of Carnot cycles that approximates the given cycle. (*c*) *a* and *b* are two arbitrary points on the cycle and 1 and 2 are reversible paths connecting them.

We can write, then, for the isothermal-adiabatic sequence of lines in Fig. 25-9$b$,

$$\sum \frac{Q}{T} = 0,$$

or, in the limit of infinitesimal temperature differences between the isotherms of Fig. 25-9$b$,

$$\oint \frac{dQ}{T} = 0, \tag{25-8}$$

in which $\oint$ indicates that the integral is evaluated for a complete traversal of the cycle, starting (and ending) at any arbitrary point of the cycle.

If the integral of a quantity around any closed path is zero, that quantity is called a state variable, that is, it has a value that is characteristic only of the state of the system, regardless of how that state was arrived at. We call the variable in this case the *entropy* $S$ and we have, from Eq. 25-8,

$$dS = \frac{dQ}{T} \quad \text{and} \quad \oint dS = 0. \tag{25-9}$$

Common units for entropy are joules/K° or cal/K°.

Gravitational potential energy $U_g$, internal energy $U$, pressure $p$, and temperature $T$ are other state variables and equations of the form $\oint dX = 0$ hold for each of them, where for $X$ we substitute the appropriate symbol. Heat $Q$ and work $W$ are *not* state variables and we know that, in general, $\oint dQ \neq 0$ and $\oint dW \neq 0$, as the student can easily show for the special case of a Carnot cycle.

The property of a state variable expressed by $\oint dX = 0$ can also be expressed by saying that $\int dX$ between any two equilibrium states has the same value for all (reversible) paths connecting those states. Let us prove this for the state variable called entropy. We can write Eq. 25-9 (see Fig. 25-9$c$) as

$${}_1\int_a^b dS + {}_2\int_b^a dS = 0 \tag{25-10}$$

where $a$ and $b$ are arbitrary points and 1 and 2 describe the paths connecting these points. Since the cycle is reversible, we can write Eq. 25-10 as

$${}_1\int_a^b dS - {}_2\int_a^b dS = 0$$

or

$${}_1\int_a^b dS = {}_2\int_a^b dS \tag{25-11}$$

In Eq. 25-11 we have simply decided to traverse path 2 in the opposite direction, that is, from $a$ to $b$ rather than from $b$ to $a$. We do this by changing the order of the limits in the second integral of Eq. 25-10, which requires that we also change the sign of the integral, thus yielding Eq. 25–11. This latter equation tells us that the quantity $\int_a^b dS$ between any two equilibrium states of the system, such as $a$ and $b$, is independent of the path connecting those states, for 1 and 2 are quite arbitrary paths. The student will recall our almost identical discussion in Section 8–2, where we introduced the concept of a conservative force.

The change in entropy between $a$ and $b$ in Fig. 25–9$c$ is, then

$$S_b - S_a = \int_a^b dS = \int_a^b \frac{dQ}{T} \qquad \text{(reversible process)}, \qquad (25\text{-}12)$$

where the integral is evaluated over *any reversible path* connecting these two states.

## 25–8 Entropy—Irreversible Processes

In Section 25–7 we spoke only of reversible processes. However, entropy, like all state variables, depends only on the state of the system and we must be able to calculate the change in entropy for irreversible processes, provided only that they begin and end in equilibrium states. Let us consider two examples.

1. *Free Expansion.* As in Section 22–7 (see Fig. 22–14) let a gas double its volume by expanding into an evacuated enclosure. Since no work is done against the vacuum, $W = 0$ and, since the gas is enclosed by non-conducting walls, $Q = 0$. From the first law, then $\Delta U = 0$ or

$$U_i = U_f \qquad (25\text{–}13)$$

where $i$ and $f$ refer to the initial and final (equilibrium) states. If the gas is an ideal gas, then $U$ depends on temperature alone and not on the pressure or the volume so that Eq. 25–13 implies $T_i = T_f$.

The free expansion is certainly irreversible because we lose control of the environment once we turn the stopcock in Fig. 22–14. There is, however, an entropy difference $S_f - S_i$ between the initial and final equilibrium states, but we cannot calculate it from Eq. 25–12 because that relation applies only to reversible paths; if we tried to use that equation we would have the immediate difficulty that $Q = 0$ for the free expansion and—further—we would not know how to assign meaningful values of $T$ to the intermediate, nonequilibrium states.

How, then, do we calculate $S_f - S_i$ for a free expansion? We do so by finding a *reversible* path (*any* reversible path) that connects the states $i$ and $f$ and we calculate the entropy change for that path. In the free expansion a convenient reversible path (assuming an ideal gas) is an isothermal expansion from $V_i$ to $V_f$ $(= 2V_i)$. This corresponds to the isothermal expansion carried out between the points $a$ and $b$ of the Carnot cycle of Fig. 25–4. It represents quite a different set of operations from the free expansion and has in common with it *only* the fact that it connects the same set of equilibrium states, *i* and *f*. From Eq. 25–12 and Example 1 we have

$$S_f - S_i = \int_i^f \frac{dQ}{T} = \mu R \ln (V_f/V_i)$$
$$= \mu R \ln 2.$$

This is positive so that the entropy of the system *increases* in this irreversible, adiabatic process.

2. *Heat Conduction.* For another example consider two bodies that are similar in every respect except that one is at a temperature $T_1$ and the other at temperature $T_2$, where $T_1 > T_2$. If we put both objects in contact inside a box with nonconducting walls, they will eventually reach a common temperature $T_m$, approximately half-way between $T_1$ and $T_2$. Like the free expansion, the process is irreversible because we lose control of the environment once we put the two bodies in the box. Like the free expansion this process is also (irreversibly) adiabatic because no heat enters or leaves the system during the process.

To calculate the entropy change for the system during this process we must again find a *reversible* process connecting the same initial and final states and calculate the system entropy change by applying Eq. 25–12 to that process. We can do so if we imagine that we have at our disposal a heat reservoir of large heat capacity whose temperature $T$ is at our control, by turning a knob, say. We first adjust the reservoir temperature to $T_1$ and put the first (hotter) object in contact with the reservoir. We then *slowly* (reversibly) lower the reservoir temperature from $T_1$ to $T_m$, extracting heat from the hot body as we do so. The hot body *loses* entropy in this process, the amount being approximately

$$\Delta S_1 = -\frac{Q}{T_{1,m}}$$

where $T_{1,m}$ is the average of $T_1$ and $T_m$ and $Q$ is the heat extracted.

We then adjust our reservoir temperature to $T_2$ and place it in contact with the second (cooler) object. We then *slowly* (reversibly) raise the reservoir temperature from $T_2$ to $T_m$, adding heat to the cool body as we do so. The cool body *gains* entropy in this process, the amount being approximately

$$\Delta S_2 = +\frac{Q}{T_{2,m}},$$

where $T_{2,m}$ is the average of $T_2$ and $T_m$ and $Q$ is the heat added.

The two bodies are now at the same temperature $T_m$ and the system, which consists of these two bodies, is now in its final equilibrium state. The change in entropy for the complete system is

$$S_f - S_i = \Delta S_1 + \Delta S_2$$
$$= -\frac{Q}{T_{1,m}} + \frac{Q}{T_{2,m}}.$$

Since $T_{1,m} > T_{2,m}$ we have $S_f > S_i$. Again, as for the free expansion, the entropy of the system has *increased* in this irreversible, adiabatic process.

In each of these examples we must distinguish carefully between the actual (irreversible) process (free expansion or heat conduction) and the reversible process that we introduce just so that we can calculate the entropy change in the actual process. We can choose *any* reversible process, as long as it connects the same initial and final state as the actual process; all such reversible processes will yield the same entropy change because this depends only on the initial and final states and not on the process connecting them—be it reversible or irreversible.

### 25–9 Entropy and the Second Law

We are now ready to formulate the second law of thermodynamics in terms of entropy. Since this law is a generalization from experience we cannot *prove* it but can only write it down and show that our statement is in agreement with experiment and is equivalent to other formulations of the second law that we have given earlier. In this spirit we assert that the second law is: *A natural process that starts in one equilibrium state and ends in another will go in the direction that causes the entropy of the system plus environment to increase.*

Following our pattern for the zeroth law and the first law of thermodynamics (see page 561) the essence of the second law, speaking loosely, is: *There exists a useful thermodynamic variable called entrophy.* The second law also tells us how to use this variable to predict whether a particular process will occur in nature.

The two experiments of Section 25–8 (free expansion and heat conduction) are consistent with the second law. The entropy of the system *increased* in each of these irreversible processes. Note that the entropy of the environment in these two cases remains unchanged because, both being carried out in adiabatic enclosures, there was no interchange of heat with the environment. Thus, as required by our statement of the second law, the entropy of the system plus environment increased for each of these (natural) processes.

In the form that we have written it the second law applies only to irreversible processes because only such processes have a "natural direction." Indeed (see Section 25–1) the understanding of the natural directions of such processes is the main concern of the second law. Reversible processes can go equally well in either direction, however, and *for reversible processes the entropy of the system plus environment remains unchanged.* This is so because if heat $dQ$ is transferred from the environment to the system the entropy of the environment *decreases* by $dQ/T$ while that of the system *increases* by $dQ/T$, the net change for the system plus environment being zero. The fact that the process is reversible means that the environment and the system can differ in temperature by only a differential amount $dT$ when the heat transfer takes place; this is in sharp contrast to our (irreversible) heat conduction problem of the previous section, in which the temperature difference of the two bodies placed in contact was large.

Another class of processes of particular interest are adiabatic processes (reversible or irreversible); they involve no transfer of heat with the environment so that the only entropy change possible is that of the system. From our statement of the second law and from our remarks about reversible processes in the paragraph above, we conclude that

$$S_f = S_i \qquad \text{(reversible adiabatic process)}$$

and

$$S_f > S_i \qquad \text{(irreversible adiabatic process),}$$

where $S_f$ and $S_i$ are the final and initial entropies of the system.

Our statement of the second law is consistent with the Clausius statement (page 628) which declares that there is no such thing as a "perfect" refrigerator (see Fig. 25-6). If there were, the entropy of the lower temperature reservoir would *decrease* by $Q/T_2$; that of the upper temperature reservoir would *increase* by $Q/T_1$; that of the system would remain unchanged because the system traverses a cycle, returning to its starting point. Thus the net change in the entropy of the system plus environment is a *decrease*, because $T_2 < T_1$. This violates the statement of the second law that we have just given and, if we wish to retain the statement, we must conclude (with Clausius) that there is no such thing as a "perfect" refrigerator.

Our statement of the second law is also consistent with the Kelvin-Planck statement (page 628) which declares that there is no such thing as a "perfect" heat engine (see Fig. 25-5). If there were, the entropy of the reservoir at temperature $T$ would *decrease* by $Q/T$; that of the system would remain unchanged because the system traverses a cycle, returning to its starting point. Thus the net change of entropy of the system plus environment is a *decrease*. This violates the statement of the second law that we have just given and, if we wish to retain the statement, we must conclude (with Kelvin) that there is no such thing as a "perfect" heat engine.

▶ **Example 3.** Compute the entropy change of a system consisting of 1.00 kg of ice at 0° C which melts (reversibly) to water at that same temperature. The latent heat of melting is 79.6 cal/gm.

The requirement that we melt the ice *reversibly* means that we must put it in contact with a heat reservoir whose temperature exceeds 0° C by only a differential amount; if we lower the reservoir temperature until it is a differential amount below 0° C, the melted ice will begin to freeze. Since the process is reversible, we can use Eq. 25-12 to compute the entropy change of the system. The temperature remains constant at 273° K. Therefore,

$$S_{\text{water}} - S_{\text{ice}} = \int_0^Q \frac{dQ}{T} = \frac{1}{T}\int_0^Q dQ = \frac{Q}{T}.$$

But

$$Q = 10^3 \text{ gm} \times 79.6 \text{ cal/gm} = 7.96 \times 10^4 \text{ cal}$$

or

$$S_{\text{water}} - S_{\text{ice}} = \frac{7.96 \times 10^4}{273} \text{ cal/°K} = 292 \text{ cal/K°}$$
$$= 1220 \text{ joules/K°}.$$

In this example of reversible melting the entropy change of the *system plus environment* is zero, as it must be for all reversible processes. The entropy change calculated above is the increase in entropy of the *system;* there is an exactly equal decrease in entropy of the environment (−1220 joules/K°) associated with the heat that leaves the reservoir (environment), at 273° K, to melt the ice.

In practice, melting is likely to be irreversible, as when we put an ice cube in a glass of water at room temperature. This process has only one natural direction—the ice will melt. The entropy of the system plus environment will *increase* in this process as required by the second law. The (irreversible) heat conduction example of the previous section should make this understandable.

▶ **Example 4.** Calculate the entropy change that an ideal gas undergoes in a reversible isothermal expansion from a volume $V_i$ to a volume $V_f$.

From the first law

$$dU = dQ - p\,dV.$$

But $dU = 0$, since $U$ depends only on temperature for an ideal gas and the temperature is constant. Hence,

$$dQ = p\,dV$$

and

$$dS = \frac{dQ}{T} = \frac{p\,dV}{T}.$$

But

$$pV = \mu RT,$$

so that

$$dS = \mu R\frac{dV}{V}$$

and

$$S_f - S_i = \int_{V_i}^{V_f} \mu R\frac{dV}{V} = \mu R \ln\frac{V_f}{V_i}. \qquad (25\text{-}14)$$

Since $V_f > V_i$, $S_f > S_i$ and the *entropy of the gas increases.*

In order to carry out this process we must have a reservoir at temperature $T$ which is in contact with the system and supplies the heat to the gas. Hence, the *entropy of the reservoir decreases* by $\int dQ/T[= \mu R \ln (V_f/V_i)]$, so that in this process the entropy of system plus environment does not change. As in the previous example, this is characteristic of a reversible process. ◀

## 25–10 Entropy and Disorder

Freeman Dyson, in an article* "What Is Heat?" writes:

"Heat is disordered energy. So with two words the nature of heat is explained. . . . Energy can exist without disorder. For example, a flying rifle bullet or an atom of $U^{235}$ carries ordered energy. The motion of the bullet is the kind we call kinetic. When the bullet hits a steel plate and is stopped, the energy of its motion is transferred to random motions of the atoms in the bullet and the plate. This disordered energy makes itself felt in the form of heat. . . . The energy dwelling in the uranium atom is the kind we call potential; it consists of the electric forces which tend to push the constituent protons apart. When the atom fissions, the energy of motion of the flying fragments is converted by collisions into random motions of the electrons and other atoms nearby in the surrounding matter—that is to say, into heat. This conversion of potential energy into heat is the working principle of nuclear reactors.

---

* *Scientific American*, September 1954.

These two examples illustrate the general principle that energy becomes heat as soon as it is disordered. It is conversely true that disorder can exist without energy, and that disorder becomes heat as soon as it is energized. The atoms of $U^{235}$ and $U^{238}$ in a piece of ordinary uranium are mixed in a random way, but this disorder carries no energy. . . .

In order to go further it is necessary to talk quantitatively. We must measure heat precisely in terms of numbers. . . . First it is clear that to specify heat we must use at least two numbers: one to measure the quantity of energy, the other to measure the quantity of disorder. The quantity of energy is measured in terms of a practical unit called the calorie. . . . The quantity of disorder is measured in terms of the mathematical concept called entropy. . . ."

If there is a connection between disorder and entropy then disorder, like entropy, must increase in natural processes. We will try to show that this is true by showing that in the examples of Section 25-8, the free expansion and heat conduction, the disorder of the system plus environment does in fact increase. We will use reasonable qualitative concepts of disorder first and will define disorder more rigorously afterwards.

1. *Free Expansion.* In a free expansion (Section 22-7) the gas molecules confined to one-half of a box are permitted to fill the entire box. By any reasonable definition of the word disorder the system has become more disordered, in the same sense that disorder increases if the litter on one vacant lot is spread over two lots. More precisely, the disorder has increased because we have lost some of our ability to classify molecules. The statement: "The molecules are in the box" is weaker from this point of view than the statement: "The molecules are in the left half of the box;" see below.

2. *Heat Conduction.* In this example two bodies of different temperatures $T_1$ and $T_2$ come to a uniform intermediate temperature $T$ when they are placed in contact. Here again the system has become more disordered in this natural process because we have lost some of our ability to classify molecules. The statement: "All molecules in the system correspond, by way of Eq. 23-6, to temperature $T$" is weaker from this point of view than the statement: "All molecules in body $A$ correspond to temperature $T_1$ and all molecules in body $B$ correspond to temperature $T_2$."

These two examples and the two given by Dyson at the beginning of this section (the bullet striking a steel plate and the fissioning uranium nucleus in a nuclear reactor) convince us that *there is a tendency for natural processes to proceed toward a state of greater disorder.*

In statistical mechanics we give a precise meaning to disorder and we express its connection with entropy by the relation

$$S = k \ln w. \tag{25-15}$$

Here, $k$ is Boltzmann's constant, $S$ is the entropy of the system, and $w$, which we may call the *disorder parameter*, is the probability that the system will exist in the state it is in relative to all the possible states it could be in. This equation connects a thermodynamic or macroscopic quantity, the entropy, with a statistical or microscopic quantity, the probability.

Let us illustrate by computing the change in entropy of an ideal gas in an isothermal expansion. Here the number of molecules and the temperature do not change, but the volume does. The probability that a given molecule may be found in a region having a volume $V$ is proportional to $V$; that is, the greater $V$ is, the greater the chance of finding it in $V$. Hence, the probability of finding a *single* molecule in $V$ is

$$w_1 = cV$$

where $c$ is a constant. The probability of finding $N$ molecules simultaneously in the volume $V$ is the $N$-fold product of $w_1$. That is, the probability of a state consisting of $N$ molecules in a volume $V$ is

$$w = w_1{}^N = (cV)^N. \tag{25–16}$$

For example, if the probability of finding a single molecule in $V$ is $\frac{1}{2}$ (that is, there is a 50% chance of its being in $V$ and a 50% chance of its being outside $V$), the probability of finding two molecules in $V$ is $\frac{1}{4}$. There are four equally probable states here (both in; both out; one in, the other out; one out, the other in), and only one of them is a state with both molecules in $V$.

If we now combine Eq. 25–15 and Eq. 25–16 we obtain

$$S = kN\,(\ln c + \ln V).$$

Hence, the difference in entropy between a state of volume $V_f$ and a state of volume $V_i$ (temperature and number of molecules remaining constant) is

$$\begin{aligned} S_f - S_i &= kN\,(\ln c + \ln V_f) - kN\,(\ln c + \ln V_i) \\ &= kN \ln \frac{V_f}{V_i} = \frac{RN}{N_0} \ln \frac{V_f}{V_i} = \mu R \ln \frac{V_f}{V_i} \end{aligned}$$

in exact agreement with the strictly thermodynamic result of Eq. 25–14.

It is on the basis of Eq. 25–16 that we stated above that disorder increases during a free expansion; that equation yields $(cV)^N$ for the disorder parameter before expansion and $(c2V)^N$ for that parameter when the volume is doubled by the expansion.

The statistical definition of entropy, Eq. 25–15, connects the thermodynamic and the statistical mechanical pictures and enables us to put the second law of thermodynamics on a statistical basis. The direction in which natural processes take place (toward higher entropy) is determined by the laws of probability (toward a more probable state). The equilibrium state is the state of maximum entropy thermodynamically and the most probable state statistically. We have seen, however, that fluctuations may occur about an equilibrium distribution (for example, the Brownian motion). From this point of view, then, it is not absolutely certain that the entropy increases in every spontaneous process. The entropy may sometimes decrease. If we waited long enough, even the most improbable states might occur: the water in a pond suddenly freezing on a hot summer day or a local vacuum occurring suddenly in a room. Although such occurrences are possible, the probability of their happening, when computed, turns out to be incredibly small. Hence, the second law of thermodynamics shows us the most probable course of events, not the only possible ones. But its area of application is so broad and the chance of nature's contradicting it so small that it occupies the distinction of being one of the most useful and general laws in all sciences.

## QUESTIONS

**1.** What requirements should a system meet in order to be in thermodynamic equilibrium?

**2.** In the irreversible process of Fig. 25–1*a* can we calculate the work done in terms of an area on a $p$-$V$ diagram? Is any work done?

**3.** Can a given amount of mechanical energy be converted completely into heat energy? If so, give an example.

4. Can you suggest a reversible process whereby heat can be added to a system? Would adding heat by means of a Bunsen burner be a reversible process?

5. Give some examples of irreversible processes in nature.

6. Give a qualitative explanation of how frictional forces between moving surfaces produce heat energy. Does the reverse process (heat energy producing relative motion of these surfaces) occur? Can you give a plausible explanation?

7. A block returns to its initial position after dissipating mechanical energy to heat through friction. Is this process reversible thermodynamically?

8. To carry out a Carnot cycle we need not start at point $a$ in Fig. 25–4, but may equally well start at points $b$, $c$, or $d$, or indeed any intermediate point. Explain.

9. If a Carnot engine is independent of the working substance, then perhaps real engines should be similarly independent, to a certain extent. Why then, for real engines, are we so concerned to find suitable fuels such as coal, gasoline, or fissionable material? Why not use stones as a fuel?

10. Couldn't we just as well define the efficiency of an engine as $e = W/Q_2$ rather than as $e = W/Q_1$? Why don't we?

11. What factors reduce the efficiency of a heat engine from its ideal value?

12. In order to increase the efficiency of a Carnot engine most effectively, would you increase $T_1$, keeping $T_2$ constant, or would you decrease $T_2$, keeping $T_1$ constant?

13. Can a kitchen be cooled by leaving the door of an electric refrigerator open? Explain.

14. Is there a change in entropy in purely mechanical motions?

15. Two samples of a gas initially at the same temperature and pressure are compressed from a volume $V$ to a volume $(V/2)$, one isothermally, the other adiabatically. In which sample is the final pressure greater? Does the entropy of the gas change in either process?

16. Suppose we had chosen to represent the state of a system by its entropy and its absolute temperature rather than by its pressure and volume. What would a Carnot cycle look like on a $T$-$S$ diagram?

17. Consider a box containing a very small number of molecules, say five. It must sometimes happen by chance that all of these molecules find themselves in the left half of the box, the right half being completely empty. This is just the reverse of a free expansion, a process that we have declared to be *irreversible*. What is your explanation?

18. Show that the total entropy increases when work is converted into heat by friction between sliding surfaces. Describe the increase in disorder.

19. Comment on the statement "A heat engine converts disordered mechanical motion into organized mechanical motion."

20. When we put cards together in a deck or put bricks together to build a house, for example, we increase the order in the physical world. Does this violate the second law of thermodynamics? Explain.

21. A rubber band feels warmer than its surroundings immediately after it is quickly stretched; it becomes noticeably cooler when it is allowed to contract rapidly; and a rubber band supporting a load contracts on being heated. Explain these observations using the fact that the molecules of rubber consist of intertwined and cross-linked long chains of atoms in roughly random orientation.

22. Explain the statement "Cosmic rays continually *decrease the entropy of the earth* on which they fall." Does this contradict the second law of thermodynamics?

23. Discuss the following comment of Panofsky and Phillips: "From the standpoint of formal physics there is only one concept which is asymmetric in the time, namely entropy. But this makes it reasonable to assume that the second law of thermodynamics can be used to ascertain the sense of time independently in any frame of refer-

ence; that is, we shall take the positive direction of time to be that of statistically increasing disorder, or increasing entropy . . ."

## PROBLEMS

**1.** An ideal gas heat engine operates in a Carnot cycle between 227 and 127° C. It absorbs $6.0 \times 10^4$ cal at the higher temperature. How much work per cycle is this engine capable of performing?

**2.** (*a*) A Carnot engine operates between a hot reservoir at 320° K and a cold reservoir at 260° K. If it absorbs 500 joules of heat at the hot reservoir, how much work does it deliver? (*b*) If the same engine, working in reverse, functions as a refrigerator between the same two reservoirs, how much work must be supplied to remove 1000 joules of heat from the cold reservoir?

**3.** In a two-stage heat engine a quantity of heat $Q_1$ is absorbed at a temperature $T_1$, work $W_1$ is done, and a quantity of heat $Q_2$ is expelled at a lower temperature $T_2$ by the first stage. The second stage absorbs the heat expelled by the first, does work $W_2$, and expels a quantity of heat $Q_3$ at a lower temperature $T_3$. Prove that the efficiency of the combination engine is $(T_1 - T_3)/T_1$.

**4.** A combination mercury-steam turbine takes saturated mercury vapor from a boiler at 876° F and exhausts it to heat a steam boiler at 460° F. The steam turbine receives steam at this temperature and exhausts it to a condenser at 100° F. What is the maximum efficiency of the combination?

**5.** Using the equation of state of an ideal gas and the equation describing an adiabatic process for an ideal gas, show that the slope, $dp/dV$, on a $p$-$V$ diagram of an adiabatic can be written as $-\gamma p/V$ and of an isothermal can be written as $-p/V$. From these results prove that adiabatics are steeper curves than isothermals.

**6.** (*a*) Plot an exact Carnot cycle on a $p$-$V$ diagram for 1 mole of an ideal gas. Let point $a$ correspond to $p = 1.0$ atm, $T = 300$° K, and let $b$ correspond to $p = 0.5$ atm, $T = 300$° K; take the low temperature reservoir to be at 100° K. Let $\gamma = 1.5$. (*b*) Compute graphically the work done in this cycle.

**7.** In a Carnot cycle, the isothermal expansion of the gas takes place at 400° K and the isothermal compression at 300° K. During the expansion 500 cal of heat energy are transferred to the gas. Determine (*a*) the work performed by the gas during the isothermal expansion, (*b*) the heat rejected from the gas during the isothermal compression, (*c*) the work done on the gas during the isothermal compression.

**8.** (*a*) If the Carnot cycle is run backward, we have an ideal refrigerator. A quantity of heat $Q_2$ is taken in at the lower temperature $T_2$ and a quantity of heat $Q_1$ is given out at the higher temperature $T_1$. The difference is the work $W$ that must be supplied to run the refrigerator. Show that

$$W = Q_2 \frac{T_1 - T_2}{T_2}.$$

(*b*) The *coefficient of performance* $K$ of a refrigerator is defined as the ratio of the heat extracted from the cold source to the work needed to run the cycle. Show that ideally

$$K = \frac{T_2}{T_1 - T_2}.$$

In actual refrigerators $K$ has a value of 5 or 6.

**9.** In a mechanical refrigerator the low-temperature coils are at a temperature of −13° C, and the compressed gas in the condenser has a temperature of 27° C. What is the theoretical coefficient of performance?

**10.** How much work must be done to transfer 1.0 joule of heat from a reservoir at 7° C to one at 27° C by means of a refrigerator using a Carnot cycle? From one at −73° C to one at 27° C? From one at −173° C to one at 27° C? From one at −223° C to one at 27° C?

**11.** The motor in a refrigerator has a power output of 200 watts. If the freezing compartment is at 270° K and the outside air is at 300° K, assuming ideal efficiency, what is the maximum amount of heat that can be extracted from the freezing compartment in 10 min?

**12.** How is the efficiency of a reversible heat engine related to the coefficient of performance of the reversible refrigerator obtained by running the engine backward?

**13.** In a heat pump, heat $Q_2$ is extracted from the outside atmosphere at $T_2$ and a larger quantity of heat $Q_1$ is delivered to the inside of the house at $T_1$, with the performance of work $W$. (*a*) Draw a schematic diagram of a heat pump. (*b*) How does it differ in principle from a refrigerator? In practical use? (*c*) How are $Q_1$, $Q_2$, and $W$ related to one another? (*d*) Can a heat pump be reversed for use in summer? Explain. (*e*) What advantages does such a pump have over other heating devices?

**14.** In a heat pump, heat from the outdoors at −5° C is transferred to a room at 17° C, energy being supplied by an electric motor. How many joules of heat will be delivered to the room for each joule of electric energy consumed, ideally?

**15.** Suppose that we were to take as our measure of temperature $-1/T$ rather than $T$. The unit of this new measure might be the Nivlek (Kelvin spelled backwards) degree (°N). Write a sequence of temperatures in °N extending from positive to negative values of $T$. (See footnote, page 532.)

**16.** (*a*) Show that when a substance of mass $m$ having a constant specific heat $c$ is heated from $T_1$ to $T_2$ the entropy change is

$$S_2 - S_1 = mc \ln \frac{T_2}{T_1}.$$

(*b*) Does the entropy of the substance decrease on cooling? If so, does the total entropy decrease in such a process? Explain.

**17.** In a specific heat experiment 100 gm of lead ($c_p = 0.0345$ cal/gmC°) at 100° C is mixed with 200 gm of water at 20° C. Find the difference in entropy of the system at the end from its value before mixing.

**18.** Four moles of an ideal gas are caused to expand from a volume $V_1$ to a volume $V_2(= 2V_1)$. (*a*) If the expansion is isothermal at the temperature $T = 400°$ K, deduce an expression for the work done by the expanding gas. (*b*) For the isothermal expansion just described, deduce an expression for the change in entropy, if any. (*c*) If the expansion were reversibly adiabatic instead of isothermal, would the change in entropy be positive, negative, or zero?

**19.** Heat can be removed from water at 0° C and atmospheric pressure without causing the water to freeze, if done with little disturbance of the water. Suppose the water is cooled to −5.0° C before ice begins to form. What is the change in entropy per unit mass occurring during the sudden freezing that then takes place?

**20.** An 8.00-gm ice cube at −10.0° C is dropped into a thermos flask containing 100 cm$^3$ of water at 20.0° C. What is the change in entropy of the system when a final equilibrium state is reached?

**21.** A brass rod is in contact thermally with a heat reservoir at 127° C at one end and a heat reservoir at 27° C at the other end. Compute the total change in the entropy arising from the process of conduction of 1200 cal of heat through the rod. Does the entropy of the rod change in the process?

**22.** A mole of a monatomic ideal gas is taken from an initial state of pressure $p$ and volume $V$ to a final state of pressure $2p$ and volume $2V$ by two different processes.

(I) It expands isothermally until its volume is doubled, and then its pressure is increased at constant volume to the final state. (II) It is compressed isothermally until its pressure is doubled, and then its volume is increased at constant pressure to the final state.

Show the path of each process on a $p$-$V$ diagram. For each process calculate in terms of $p$ and $V$, or of $T$, (*a*) the heat absorbed by the gas in each part of the process; (*b*) the work done on the gas in each part of the process; (*c*) the change in internal energy of the gas $U_f - U_i$; (*d*) the change in entropy of the gas $S_f - S_i$.

**23.** One mole of hydrogen gas and 1.0 mole of nitrogen gas are in adjacent containers at the same pressure $p$ and temperature $T$. The pressure and temperature are such that both gases behave virtually ideally. (*a*) If the rms speed of the $H_2$ molecules is 1850 meters/sec at temperature $T$, what will the rms speed be of the $N_2$ molecules? (*b*) For which gas will a larger percentage or fraction of the molecules have speeds within ±50 meters/sec of the rms speed? (*c*) If the containers are connected so that the $H_2$ and $N_2$ mix, will the change in entropy be positive, negative, or zero?

**24.** (*a*) A body of finite mass is originally at temperature $T_2$, higher than that of a heat reservoir at a temperature $T_1$. An engine operates in infinitesimal cycles between the body and the reservoir until it lowers the temperature of the body from $T_2$ to $T_1$. Prove that the maximum work obtainable from the engine is $W_{\max} = Q - T_1(S_2 - S_1)$, where $S_1 - S_2$ is the entropy change in the body and $Q$ is the heat extracted from the body by the engine. (*b*) A body of finite mass is originally at temperature $T_1$, the same as that of a heat reservoir. A refrigerator operates in infinitesimal cycles between the body and reservoir until it lowers the temperature of the body from $T_1$ to $T_0$. Prove that the minimum amount of work which must be supplied to the refrigerator is $W_{\min} = T_1(S_1 - S_0) - Q$, where $S_0 - S_1$ is the entropy change in the body and $Q$ is the heat extracted from the body by the refrigerator.

**25.** In general, the probability $w_{12}$ of a complex event, which consists of two unrelated simple events, is equal to the product of their respective probabilities $w_1$, $w_2$. The entropy $S_{12}$ of a complex system which consists of two simple systems is just the sum of their respective entropies, $S_1$, $S_2$. Show that Eq. 25–15, which relates probability and entropy, is consistent with the additive property of entropy and the multiplicative property of probability for a complex system.