MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

SUMY STATE UNIVERSITY

A. Yunda

Lecture notes

in **Medical and Biological Physics**

In two parts

Part 1

Sumy
Sumy State University
2015

MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

SUMY STATE UNIVERSITY

A. Yunda

Lecture notes
in **Medical and Biological Physics**

In two parts

Part 1

APPROVED
at a session of Complex
Systems Modeling Department
as a lecture notes on discipline
“Medical and Biological Physics”.
Minutes № 6 from 04.02.2015

Sumy
Sumy State University
2015

Lecture notes in Medical and Biological Physics : in two parts / compiler A. Yunda. – Sumy : Sumy State University, 2015. – Part 1. – 136 p.

Complex Systems Modeling Department

Contents

Lecture 1

Physical Properties of Fluids

Matter most commonly exists as a *solid*, *liquid*, or *gas*; these states are known as the three common phases of matter. Solids have a definite shape and a specific volume, liquids have a definite volume but their shape changes depending on the container in which they are held, and gases have neither a definite shape nor a specific volume as their molecules move to fill the container in which they are held. Liquids and gases are considered to be *fluids* because they yield to shearing forces, whereas solids resist them. Note that the extent to which fluids yield to shearing forces (and hence flow easily and quickly) depends on a quantity called the *viscosity*. Liquids deform easily when stressed and do not spring back to their original shape once the force is removed because the atoms are free to slide about and change neighbors – that is, they flow (so they are a type of fluid) with the molecules held together by their mutual attraction. Because the atoms are closely packed, liquids, like solids, resist compression.

1.1. Density

Density is an important characteristic of substances. Density is the mass per unit volume of a substance or object. In equation form, density is defined as

$$\rho = \frac{m}{V}, \tag{1.1}$$

where the Greek letter ρ (rho) is the symbol for density, m is the mass, and V is the volume occupied by the substance. The SI unit of density is kg/m^3.

1.2. Pressure

You have no doubt heard the word "pressure" being used in relation to blood (high or low blood pressure) and in relation to the weather (high- and low-pressure weather systems). These are only two of many examples of pressures in fluids. Pressure P is defined as the force divided by the area perpendicular to the force over which the force is applied, or

$$P = \frac{F}{A}, \tag{1.2}$$

where F is a force applied to an area A that is perpendicular to the force. The SI unit for pressure is the pascal, where $1\,Pa = 1\,N/m^2$.

In addition to the pascal, there are many other units for pressure that are in common use. In meteorology, atmospheric pressure is often described in units of millibar (mb), where $100\,mb = \ 1 \times 10^5\ Pa$. Pounds per square inch lb/in^2 or psi is still sometimes used as a measure of tire pressure, and millimeters of mercury ($mm\ Hg$) is still often used in the measurement of blood pressure. Pressure is defined for all states of matter but is particularly important when discussing fluids.

Table 1.1 – Densities of Various Substances

Substance	ρ, $\times 10^3\ kg/m^3$	**Substance**	ρ, $\times 10^3\ kg/m^3$	**Substance**	ρ, kg/m^3
Solids		**Liquids**		**Gases**	
Aluminum	2.7	Water (4°C)	1.000	Air	1.29
Brass	8.44	Blood	1.05	Carbon dioxide	1.98
Copper (average)	8.8	Sea water	1.025	Carbon monoxide	1.25
Gold	19.32	Mercury	13.6	Hydrogen	0.090
Iron or steel	7.8	Ethyl alcohol	0.79	Helium	0.18
Lead	11.3	Petrol	0.68	Methane	0.72
Polystyrene	0.10	Glycerin	1.26	Nitrogen	1.25
Tungsten	19.30	Olive oil	0.92	Nitrous oxide	1.98
Uranium	18.70			Oxygen	1.43
Concrete	2.30-3.0			Steam (100°C)	40.60
Cork	0.24				
Glass	2.6				
Granite	2.7				
Earth's crust	3.3				
Wood	0.3-0.9				
Ice (0°C)	0.917				
Bone	1.7-2.0				

Atmospheric pressure is another example of pressure due to the weight of a fluid, in this case due to the weight of air above a given height, the average pressure at sea level is given by the standard atmospheric pressure P_{atm}, measured to be $1\,atmosphere\,(atm) = P_{atm} = 1.01\times10^5\ N/m^2 = 101\ kPa$.

1.3. Pascal's Principle

Can pressure be increased in a fluid by pushing directly on the fluid? Yes, but it is much easier if the fluid is enclosed. The heart, for example, increases blood pressure by pushing directly on the blood in an enclosed system (valves closed in a chamber). If you try to push on a fluid in an open system, such as a river, the fluid flows away. An enclosed fluid cannot flow away, and so pressure is more easily increased by an applied force. What happens to a pressure in an enclosed fluid? Since atoms in a fluid are free to move about, they transmit the pressure to all parts of the fluid and to the walls of the container. Remarkably, the pressure is transmitted undiminished. This phenomenon is called *Pascal's principle*, because it was first clearly stated by the French philosopher and scientist *Blaise Pascal* (1623-1662): A change in pressure applied to an enclosed fluid is transmitted undiminished to all portions of the fluid and to the walls of its container. *Pascal's principle* implies that the total pressure in a fluid is the sum of the pressures from different sources.

1.4. Pressure Measurement

Gauge pressure is the pressure relative to atmospheric pressure. Gauge pressure is positive for pressures above atmospheric pressure, and negative for pressures below it.

The *total pressure*, or *absolute pressure*, is the sum of gauge pressure and atmospheric pressure: $P_{abs} = P_g + P_{atm}$, where P_{abs} is absolute pressure, P_g is gauge pressure.

There is a host of devices for measuring pressure, ranging from tire gauges to blood pressure cuffs. The undiminished transmission of pressure through a fluid allows precise remote sensing of pressures. Remote sensing is often more convenient than putting a measuring device into a system, such as a person's artery. There are many types of mechanical pressure gauges in use today. In all mechanical pressure gauges, pressure results in a force that is converted (or transduced) into some type of readout.

Mercury manometers are often used to measure arterial blood pressure. An inflatable cuff is placed on the upper arm. By squeezing the bulb, the person making the measurement exerts pressure, which is transmitted undiminished to both the main artery in the arm and the manometer. When this applied pressure exceeds blood pressure, blood flow below the cuff is cut off. The person making the measurement then slowly lowers the applied pressure and listens for blood flow to resume. Blood pressure pulsates because of the pumping action of the heart, reaching a maximum, called *systolic pressure*, and a minimum, called *diastolic pressure*, with each heartbeat. Systolic pressure is measured by noting the value of h – height of the mercury column, when blood flow first begins as cuff pressure is lowered. Diastolic pressure is measured by noting h when blood flows without interruption. The typical blood pressure of a young adult raises the mercury to a height of $120\,mm$ at systolic and $80\,mm$ at diastolic. This is commonly quoted as 120 over 80, or 120/80. The first pressure is representative of the maximum output of the heart; the second is due to the elasticity of the arteries in maintaining the pressure between beats. The density of the mercury fluid in the manometer is 13.6 times greater than water, so the height of the fluid will be 1/13.6 of that in a water manometer. This reduced height can make measurements difficult, so mercury manometers are used to measure larger pressures, such as blood pressure. The density of mercury is such that $1.0\,mm\,Hg = 133\,Pa$.

1.5. Archimedes' Principle

According to Archimedes' Principle, the buoyant force on an object equals the weight of the fluid it displaces. In equation form, Archimedes' principle is

$$F_B = w_{fl}, \tag{1.3}$$

where F_B is the buoyant force and $w_{fl} = \rho_{fl} g V_o$ is the weight of the fluid displaced by the object (ρ_{fl} – density of the fluid, $g = 9.8\,m/s^2$ – gravity, V_o – the volume of fluid displaced by the object).

Density plays a crucial role in Archimedes' principle. The average density of an object is what ultimately determines whether it floats. If its average density is less than that of the surrounding fluid, it will float. This is because the fluid, having a higher density, contains more mass and hence more weight in the same volume. The buoyant force, which equals the weight of the fluid displaced, is thus greater than the weight of the object. Likewise, an object denser than the fluid will sink.

1.6. Surface Tension

Attractive forces between molecules of the same type are called *cohesive forces*. Liquids, for example, can be held in open containers because cohesive forces hold the molecules together. Attractive forces between molecules of different types are called *adhesive forces*.

Forces between atoms and molecules underlie the macroscopic effect called *surface tension.* These attractive forces pull the molecules closer together and tend to minimize the surface area. Molecules on the surface are pulled inward by cohesive forces, reducing the surface area. Molecules inside the liquid experience zero net force, since they have neighbors on all sides.

Surface tension is proportional to the strength of the cohesive force, which varies with the type of liquid. Surface tension σ is defined to be the force F per unit length L exerted by a stretched liquid membrane:

$$\sigma = \frac{F}{L}. \tag{1.4}$$

Surface tension is the reason why liquids form bubbles and droplets. The inward surface tension force causes bubbles to be approximately spherical and raises the pressure of the gas trapped inside relative to atmospheric pressure outside. It can be shown that the gauge pressure P inside a spherical bubble is given by

$$P = \frac{4\sigma}{r}, \tag{1.5}$$

where r is the radius of the bubble.

1.7. Adhesion and Capillary Action

Why is it that water beads up on a waxed car but does not on bare paint? The answer is that the adhesive forces between water and wax are much smaller than those between water and paint. Competition between the forces of adhesion and cohesion are important in the macroscopic behavior of liquids. An important factor in studying the roles of these two forces is the angle θ between the tangent to the liquid surface and the surface. (See Figure 1.1) The contact angle θ is directly related to the relative

Table 1.2 – Surface Tension of Some Liquids

Liquid	**Surface tension** σ, N/m
Water at 0°C	0.0756
Water at 20°C	0.0728
Water at 100°C	0.0589
Soapy water (typical)	0.0370
Ethil alcohol	0.0223
Glycerin	0.0631
Mercury	0.465
Olive oil	0.032
Tissue fluids (typical)	0.050
Blood, whole at 37°C	0.058
Blood plasma at 37°C	0.073
Gold at 1070°C	1.000
Oxygen at -193°C	0.0157
Helium at -269°C	0.00012

strength of the cohesive and adhesive forces. The larger the strength of the cohesive force relative to the adhesive force, the larger θ is, and the more the liquid tends to form a droplet. The smaller θ is, the smaller the relative strength, so that the adhesive force is able to flatten the drop.

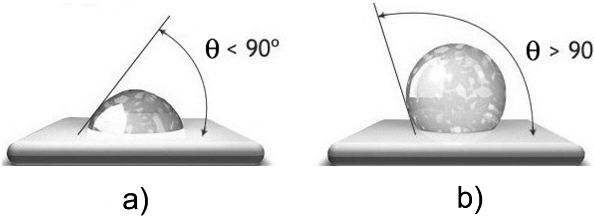

Figure 1.1 – a) Water beads on bare paint are flattened considerably because the adhesive forces between water and paint are strong, overcoming surface tension. b) Water forms beads on the waxed surface because the cohesive forces responsible for surface tension are larger than the adhesive forces, which tend to flatten the drop

One important phenomenon related to the relative strength of cohesive and adhesive forces is *capillary action* – the tendency of a fluid to be raised or suppressed in a narrow tube, or *capillary tube*. This action causes blood to be drawn into a small-diameter tube when the tube touches a drop.

If a capillary tube is placed vertically into a liquid, as shown in Figure 1.2, capillary action will raise or suppress the liquid inside the tube depending on the combination of substances. The actual effect depends on the relative strength of the cohesive and adhesive forces and, thus, the contact angle θ. If θ is less than 90°, then the fluid will be raised; if θ is greater than 90°, it will be suppressed. Mercury, for example, has a very large surface tension and a large contact angle with glass. When placed in a tube, the surface of a column of mercury curves downward, somewhat like a drop. The curved surface of a fluid in a tube is called a *meniscus.* The tendency of surface tension is always to reduce the surface area. Surface tension thus flattens the curved liquid surface in a capillary tube. This results in a downward force in mercury and an upward force in water.

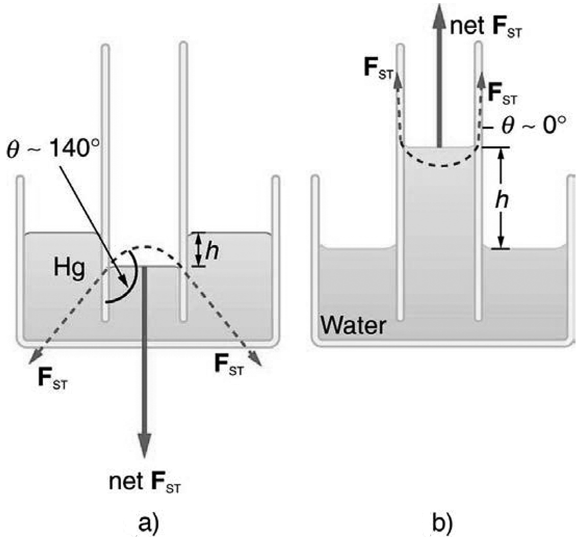

Figure 1.2 – a) Mercury is suppressed in a glass tube because its contact angle is greater than 90°. b) Water is raised in a glass tube because its contact angle is nearly 90°. Here F_{ST} is surface tension

Capillary action can move liquids horizontally over very large distances, but the height to which it can raise or suppress a liquid in a tube is limited by its weight. It can be shown that this height h is given by

$$h = \frac{2\sigma \cos(\theta)}{\rho g r}. \tag{1.6}$$

The height is directly proportional to the surface tension σ, which is its direct cause. Furthermore, the height is inversely proportional to tube radius – the smaller the radius r, the higher the fluid can be raised, since a smaller tube holds less mass. The height is also inversely proportional to fluid density ρ, since a larger density means a greater mass in the same volume.

1.8. Pressures in the Body

Next to taking a person's temperature and weight, measuring blood pressure is the most common of all medical examinations. Control of high blood pressure is largely responsible for the significant decreases in heart attack and stroke fatalities achieved in the last three decades. The pressures in various parts of the body can be measured and often provide valuable medical indicators. In this section, we consider a few examples together with some of the physics that accompany them.

Blood Pressure

Common arterial blood pressure measurements typically produce values of 120 $mm\ Hg$ and 80 $mm\ Hg$, respectively, for systolic and diastolic pressures. Both pressures have health implications. When systolic pressure is chronically high, the risk of stroke and heart attack is increased. If, however, it is too low, fainting is a problem. Systolic pressure increases dramatically during physical exercises to increase blood flow and returns to normal afterward. This change produces no ill effects and, in fact, may be beneficial to the tone of the circulatory system. Diastolic pressure can be an indicator of fluid balance. When

Table 1.3 – Typical Pressures in Humans

Body system	**Gauge pressure,** *mm Hg*
Blood pressure in large arteries (resting)	
Maximum (systolic)	100-140
Minimum (diastolic)	60-90
Blood pressure in large veins	4-15
Eye	12-24
Brain and spinal fluid (lying down)	5-12
Bladder	
While filling	0-25
When full	100-150
Chest cavity between lungs and ribs	-8 to -4
Inside lungs	-2 to $+3$
Digestive tract	
Esophagus	-2
Stomach	0-20
Intestines	10-20
Middle ear	< 1

low, it may indicate that a person is hemorrhaging internally and needs a transfusion. Conversely, high diastolic pressure indicates a ballooning of the blood vessels, which may be due to the transfusion of too much fluid into the circulatory system. High diastolic pressure is also an indication that blood vessels are not dilating properly to pass blood through. This can seriously strain the heart in its attempt to pump blood. Blood leaves the heart at about 120 $mm\ Hg$, but its pressure continues to decrease (to almost 0) as it goes from the aorta to smaller arteries to small veins (see Figure 1.3). The pressure differences in the circulation system are caused by blood flow through the system as well as the position of the person. For a person standing up, the pressure in the feet will be larger than at the heart due to the weight of the blood ($P = h\rho g$). If we assume that the distance between the heart and the feet of a person in an upright position is $1.4\,m$, then the increase in pressure in the feet relative to that in the heart (for a static column of blood) is given by

$$\begin{aligned} \Delta P = \Delta h\rho g = (1.4\,m)(1050\,kg/m^3)(9.80\,m/s^2) = \\ = 1.4 \times 10^4\ Pa = 108\,mmHg. \end{aligned} \tag{1.7}$$

Standing a long time can lead to an accumulation of blood in the legs and swelling. This is the reason why soldiers who are required to stand still for long periods of time have been known to faint. Elastic bandages around the calf can help prevent this accumulation and can also help provide increased pressure to enable the veins to send blood back up to the heart. For similar reasons, doctors recommend tight stockings for long-haul flights.

Blood pressure may also be measured in the major veins, the heart chambers, arteries to the brain, and the lungs. But these pressures are usually monitored only during surgery or for patients in intensive care since the measurements are invasive. To obtain these pressure measurements, qualified health care workers thread

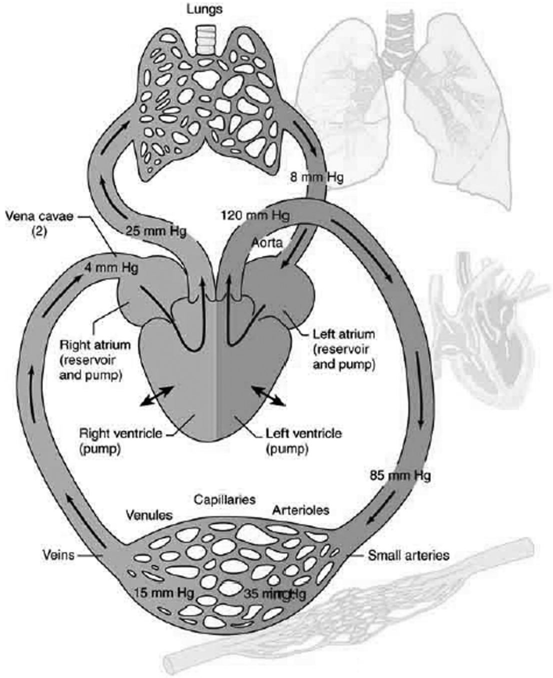

Figure 1.3 – Schematic of the circulatory system showing typical pressures. The two pumps in the heart increase pressure and that pressure is reduced as the blood flows through the body

thin tubes, called catheters, into appropriate locations to transmit pressures to external measuring devices. The heart consists of two pumps: the right side forcing blood through the lungs and the left side causing blood to flow through the rest of the body (Figure 1.3). Right-heart failure, for example, results in a rise in the pressure in the vena cavae and a drop in pressure in the arteries to the lungs. Left-heart failure results in a rise in the blood pressure entering the left side of the heart and a drop in aortal pressure.

Pressure in the Eye

The shape of the eye is maintained by fluid pressure, called intraocular pressure, which is normally in the range of 12.0 to 24.0 $mm\ Hg$. When the circulation of fluid in the eye is blocked, it can lead to pressure buildup, a condition called *glaucoma*. The net pressure can become as great as 85.0 $mm\ Hg$, an abnormally large pressure that can permanently damage the optic nerve.

Pressure Associated with the Lungs

The pressure inside the lungs increases and decreases with each breath. The pressure drops to below atmospheric pressure (negative gauge pressure) when you inhale, causing air to flow into the lungs. It increases above atmospheric pressure (positive gauge pressure) when you exhale, forcing air out. Lung pressure is controlled by several mechanisms. Muscle action in the diaphragm and rib cage is necessary for inhalation; this muscle action increases the volume of the lungs thereby reducing the pressure within them. Surface tension in the alveoli creates a positive pressure opposing inhalation. You can exhale without muscle action by letting surface tension in the alveoli create its own positive pressure. Muscle action can add to this positive pressure to produce forced exhalation, such as when you blow up a balloon, blow out a candle, or cough. The lungs, in fact, would collapse due to the surface tension in the alveoli if they were not attached to the inside of the chest wall by liquid adhesion. The gauge pressure in the liquid attaching the lungs to the inside of the chest wall is thus negative, ranging from -4 to $-8\,mm\,Hg$ during exhalation and inhalation, respectively. If air is allowed to enter the chest cavity, it breaks the attachment, and one or both lungs may collapse. Suction is applied to the chest cavity of surgery patients and trauma victims to reestablish negative pressure and inflate the lungs.

Spinal Column and Skull

Normally, there is a 5 to 12 $mm\ Hg$ pressure in the fluid surrounding the brain and filling the spinal column. This cerebrospinal fluid serves many purposes, one of which is to supply flotation to the brain. The buoyant force supplied by the fluid nearly equals the weight of the brain, since their densities are nearly equal. If there is a loss of fluid, the brain rests on the inside of the skull, causing severe headaches, constricted blood flow, and serious damage. Spinal fluid pressure is measured by means of a needle inserted between vertebrae that transmits the pressure to a suitable measuring device.

Bladder Pressure

This bodily pressure is one of which we are often aware. In fact, there is a relationship between our awareness of this pressure and a subsequent increase in it. Bladder pressure climbs steadily from zero to about 25 $mm\ Hg$ as the bladder fills to its normal capacity of 500 cm^3. This pressure triggers the micturition reflex, which stimulates the feeling of needing to urinate. Furthermore, it also causes muscles around the bladder to contract, raising the pressure to over 100 $mm\ Hg$, accentuating the sensation. Coughing, straining, tensing in cold weather, wearing tight clothes, and experiencing simple nervous tension – all this can increase bladder pressure and trigger this reflex. So can the weight of a pregnant woman's fetus, especially if it is kicking vigorously or pushing down with its head! Bladder pressure can be measured by a catheter or by inserting a needle through the bladder wall and transmitting the pressure to an appropriate measuring device. One hazard of high bladder pressure (sometimes created by an obstruction) is that such pressure can force urine back into the kidneys, causing potentially severe damage.

Pressures in the Skeletal System

These pressures are the largest in the body, due both to the high values of initial force and the small areas to which this force is applied, such as in the joints. For example, when a person lifts an object improperly, a force of 5000 N may be created between vertebrae in the spine, and this may be applied to an area as small as 10 cm^2. The created pressure is $P = = F/A = (5000\,N)/(10^{-3}\,m^2) = 5.0\times10^6\,N/m^2$ or about 50 atm! This pressure can damage both the spinal discs (the cartilage between vertebrae), and the bony vertebrae themselves. Even under normal circumstances, forces between vertebrae in the spine are large enough to create pressures of several atmospheres. Most causes of excessive pressure in the skeletal system can be avoided by lifting properly and avoiding extreme physical activity.

Other Pressures in the Body

There are many other interesting and medically significant pressures in the body. For example, pressure caused by various muscle actions drives food and waste through the digestive system. Stomach pressure behaves much like bladder pressure and is tied to the sensation of hunger. Pressure in the relaxed esophagus is normally negative because pressure in the chest cavity is normally negative. Positive pressure in the stomach may thus force acid into the esophagus, causing "heartburn". Pressure in the middle ear can result in significant force on the eardrum if it differs greatly from atmospheric pressure while scuba diving. The decrease in external pressure is also noticeable during plane flights (due to a decrease in the weight of air above relative to that at the Earth's surface). The Eustachian tubes connect the middle ear to the throat and allow us to equalize pressure in the middle ear to avoid an imbalance of force on the eardrum. Many pressures in the human body are associated with the flow of fluids.

Lecture 2

Fluid Dynamics

We have dealt with many situations in which fluids are static. But by their very definition, fluids flow. Examples come easily – a column of smoke rises from a camp fire, water streams from a fire hose, blood courses through your veins. The physics of fluids in motion – *fluid dynamics.*

2.1. Flow Rate

Flow rate Q is defined to be the volume of fluid passing by some location through an area during a period of time, as seen in Figure 2.1. In symbols, this can be written as

$$Q = \frac{V}{t}, \tag{2.1}$$

where V is the volume and t is the elapsed time. The SI unit for flow rate is m^3/s, but a number of other units for Q are in common use. For example, the heart of a resting adult pumps blood at a rate of 5.00 liters per minute (L/min).

Flow rate and velocity are related, but quite different physical quantities. The precise relationship between flow rate Q and velocity v is

$$Q = Av, \tag{2.2}$$

where A is the cross-sectional area and v is the average velocity.

Figure 2.1 illustrates how this relationship can be obtained. The shaded cylinder has a volume

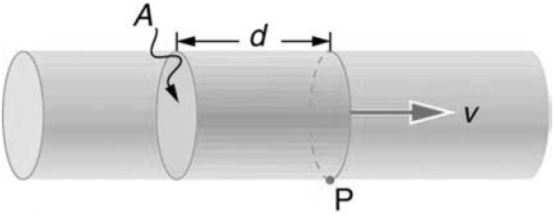

Figure 2.1 – Flow rate is the volume of fluid per unit time flowing past a point through the area A. Here the shaded cylinder of fluid flows past point P in a uniform pipe in time t

$$V = Ad,$$

which flows past the point P in a time t . Dividing both sides of this relationship by t gives

$$\frac{V}{t} = \frac{Ad}{t}.$$

We note that $Q = V/t$ and the average speed is $v = d/t$. Thus the equation becomes $Q = Av$.

Figure 2.2 shows an incompressible fluid flowing along a pipe of decreasing radius. Because the fluid is incompressible, the same amount of fluid must flow past any point in the tube in a given time to ensure continuity of flow. In this case, because the cross-sectional area of the pipe decreases, the velocity must necessarily increase. This logic can be extended to say that the flow rate must be the same at all points along the pipe. In particular, for points 1 and 2,

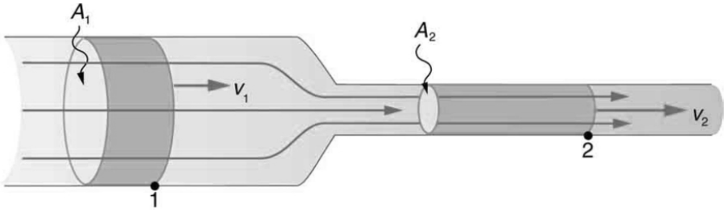

Figure 2.2 – An incompressible fluid flowing along a pipe of decreasing radius

$$\begin{cases} Q_1 = Q_2 \\ A_1 v_1 = A_2 v_2 \end{cases} \tag{2.3}$$

This is called *the equation of continuity* and is valid for any incompressible fluid. Since liquids are essentially incompressible, the equation of continuity is valid for all liquids. However, gases are compressible, and so the equation must be applied with caution to gases if they are subjected to compression or expansion.

In many situations, including the cardiovascular system, branching of the flow occurs. The blood is pumped from the heart into arteries that subdivide into smaller arteries (arterioles) which branch into very fine vessels called capillaries. In this situation, continuity of flow is maintained, but it is the sum of the flow rates in each of the branches in any portion along the tube that is maintained. The equation of continuity in a more general form becomes

$$n_1 A_1 v_1 = n_2 A_2 v_2, \tag{2.4}$$

where n_1 and n_2 are the number of branches in each of the sections along the tube.

2.2. Bernoulli's Equation

When a fluid flows into a narrower channel, its speed increases. That means its kinetic energy also increases. Where does that change in kinetic energy come from? The increased kinetic energy comes from the net work done on the fluid to push it into the channel and the work done on the fluid by the gravitational force if the fluid changes its vertical position. Recall the work-energy theorem,

$$W_{net} = \frac{mv^2}{2} - \frac{mv_0^2}{2}. \tag{2.5}$$

There is a pressure difference when the channel narrows. This pressure difference results in a net force on the fluid: recall that pressure times area equals force. The net work done increases the fluid's kinetic energy. As a result, the pressure will drop in a rapidly-moving fluid, whether or not the fluid is confined to a tube.

The relationship between pressure and velocity in fluids is described quantitatively by *Bernoulli's equation*, named after its discoverer, the Swiss scientist Daniel Bernoulli (1700–1782). Bernoulli's equation states that for an incompressible, frictionless fluid, the following sum is constant:

$$P + \frac{\rho v^2}{2} + \rho g h = constant, \tag{2.6}$$

where P is the absolute pressure, ρ is the fluid density, v is the velocity of the fluid, h is the height above some reference point, and g is the acceleration due to gravity. If we follow a small volume of fluid along its path, various quantities in the sum may change, but the total remains constant. Let the subscripts 1 and 2 refer to any two points along the path that the portion of the fluid follows; Bernoulli's equation becomes

$$P_1 + \frac{\rho v_1^2}{2} + \rho g h_1 = P_2 + \frac{\rho v_2^2}{2} + \rho g h_2. \tag{2.7}$$

Bernoulli's equation is a form of the conservation of energy principle. Note that the second and third terms are the kinetic and potential energy with m replaced by ρ.

The general form of Bernoulli's equation has three terms in it, and it is broadly applicable. To understand it better, we will look at a number of specific situations that simplify and illustrate its use and meaning.

Bernoulli's Equation for Static Fluids

Let us first consider the very simple situation where the fluid is static – that is, $v_1 = v_2 = 0$. Bernoulli's equation in that case is

$$P_1 + \rho g h_1 = P_2 + \rho g h_2. \tag{2.8}$$

We can further simplify the equation by taking $h_2 = 0$ (we can always choose some height to be zero, just as we often have done for other situations involving the gravitational force, and take all other heights to be relative to this). In that case, we get

$$P_2 = P_1 + \rho g h_1. \tag{2.9}$$

This equation tells us that, in static fluids, pressure increases with depth.

Bernoulli's Principle

Another important situation is one in which the fluid moves but its depth is constant - that is, $h_1 = h_2$. Under that condition, Bernoulli's equation becomes

$$P_1 + \frac{\rho v_1^2}{2} = P_2 + \frac{\rho v_2^2}{2}. \tag{2.10}$$

Situations in which fluid flows at a constant depth are so important that this equation is often called *Bernoulli's principle*. It is Bernoulli's equation for fluids at constant depth. Bernoulli's principle – pressure drops as speed increases in a moving fluid. For example, if v_2 is greater than v_1 in the equation, then P_2 must be less than P_1 for the equality to hold.

2.3. Laminar Flow and Viscosity

In the previous sections we have considered ideal fluids with little or no viscosity. In this section, we will investigate what factors, including viscosity, affect the rate of fluid flow.

The precise definition of *viscosity* is based on *laminar*, or nonturbulent, flow. First, we define viscosity, then, we need to define laminar flow and turbulent flow. Figure 2.3 shows both types of flow. Laminar flow is characterized by the smooth flow of the fluid in layers that do not mix. *Turbulent* flow, or turbulence, is characterized by eddies and swirls that mix layers of fluid together.

Figure 2.4 shows how viscosity is measured for a fluid. Two parallel plates have the specific fluid between them. The bottom plate is held fixed, while the top plate is moved to the right, dragging fluid with it. The layer (or lamina) of fluid in contact with either plate does not move relative to the plate, and so the top layer moves at v while the bottom layer remains at rest. Each successive layer from the top down exerts a force on the one below

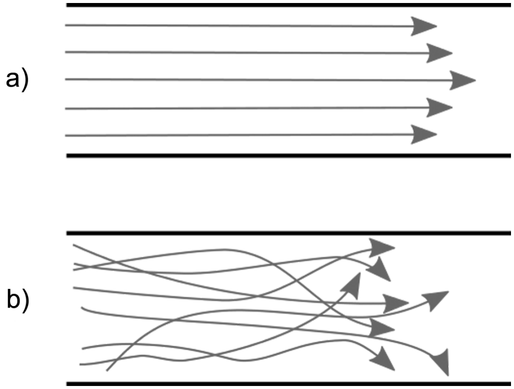

Figure 2.3 – a) Laminar flow occurs in layers without mixing. b) Turbulent flow mixes the fluid

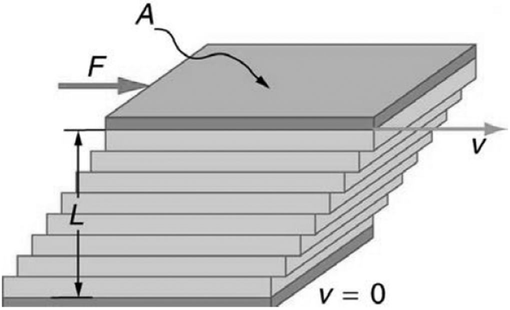

Figure 2.4 – The graphic shows laminar flow of fluid between two plates of area A. The bottom plate is fixed. When the top plate is pushed to the right, it drags the fluid along with it

it, trying to drag it along, producing a continuous variation in speed from v to 0 as shown. Care is taken to insure that the flow is laminar, that is, the layers do not mix. The motion in Figure 2.4 is like a continuous shearing motion. Fluids have zero shear strength, but the rate at which they are sheared is related to the same geometrical factors A and L as is shear deformation for solids.

A force F is required to keep the top plate in Figure 2.4 moving at a constant velocity v, and experiments have shown that this force depends on four factors. First, F is directly proportional to v (until the speed is so high that turbulence occurs – then a much larger force is needed, and it has a more complicated dependence on v). Second, F is proportional to the area A of the plate. This relationship seems reasonable since A is directly proportional to the amount of fluid being moved. Third, F is inversely proportional to the distance between the plates L. This relationship is also reasonable; L is like a lever arm, and the greater the lever arm, the less the force is needed. Fourth, F is directly proportional to the coefficient of viscosity, η. The greater the viscosity, the greater the force required. These dependencies are combined into the equation

$$F = \eta \frac{vA}{L}, \tag{2.11}$$

which gives us a working definition of fluid viscosity η.

The SI unit of viscosity is $(N/m^2)s$ or $Pa \cdot s$. Table 2.1 lists the coefficients of viscosity for various fluids.

2.4. Poiseuille's Law

In fact, there is a very simple relationship between horizontal flow and pressure. Flow rate Q is in the direction from high to low pressure. The greater the pressure differential between two points, the greater the flow rate. This relationship can be stated as

$$Q = \frac{P_2 - P_1}{R}, \tag{2.12}$$

Table 2.1 – Coefficients of Viscosity of Various Fluids

Fluid	**Temperature**, $^{\circ}C$	**Viscosity** η, $mPa \cdot s$
Gases		
Air	0	0.0171
	20	0.0181
Ammonia	20	0.00974
Carbon dioxide	20	0.0147
Helium	20	0.0196
Hydrogen	07	0.0090
Mercury	20	0.0450
Oxygen	20	0.0203
Steam	100	0.0130
Liquids		
Water	20	1.002
	37	0.6947
	100	0.282
Whole blood	20	3.015
	37	2.084
Blood plasma	20	1.810
	37	1.257
Ethyl alcohol	20	1.20
Methanol	20	0.584
Oil (olive)	20	138
Glycerin	20	1500
Honey	20	2000-10000
Milk	20	3.0

where P_1 and P_2 are the pressures at two points, such as at either end of a tube, and R is the resistance to flow. The resistance R includes everything, except pressure, that affects flow rate. For example, R is greater for a long tube than for a short one. The greater the viscosity of a fluid, the greater the value of R. Turbulence greatly increases R, whereas increasing the diameter of a tube decreases R. If viscosity is zero, the fluid is frictionless and the resistance to flow is also zero. Comparing frictionless flow in a tube to viscous flow, as in Figure 2.5, we see that for a viscous fluid, speed is the greatest at midstream because of drag at the boundaries.

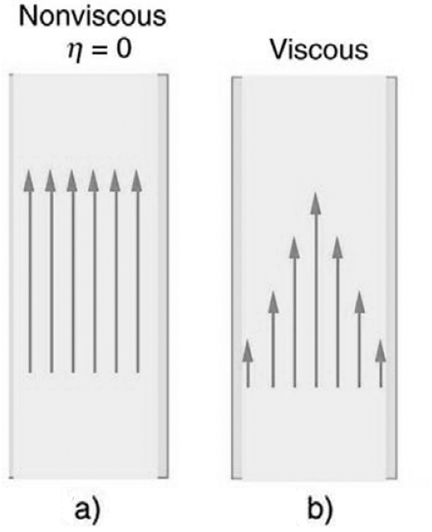

Figure 2.5 – a) If fluid flow in a tube has negligible resistance, the speed is the same all across the tube. b) When a viscous fluid flows through a tube, its speed at the walls is zero, increasing steadily to its maximum at the center of the tube

The resistance R to laminar flow of an incompressible fluid having viscosity η through a horizontal tube of uniform radius r and length l, such as the one in Figure 2.6, is given by

$$R = \frac{8\eta l}{\pi r^4}. \tag{2.13}$$

This equation is called *Poiseuille's law* for resistance after the French scientist J. L. Poiseuille (1799–1869), who derived it in an attempt to understand the flow of blood, an often turbulent fluid.

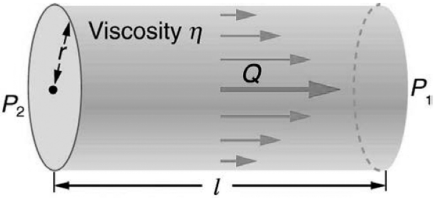

Figure 2.6 – Poiseuille's law applies to laminar flow of an incompressible fluid of viscosity η through a tube of length l and radius r

The taken together, (2.12) and (2.13) give the following expression for flow rate:

$$Q = \frac{(P_2 - P_1)\pi r^4}{8\eta l}. \tag{2.14}$$

This equation describes laminar flow through a tube. It is sometimes called Poiseuille's law for laminar flow, or simply *Poiseuille's law.*

2.5. Reynolds number

An indicator called the *Reynolds number* R_e can reveal whether the flow is laminar or turbulent. For the flow in a tube of uniform diameter, the Reynolds number is defined as

$$R_e = \frac{\rho v d}{\eta}, \tag{2.15}$$

where ρ is the fluid density, v its speed, η its viscosity, and d – the tube diameter. The Reynolds number is a unitless quantity. Experiments have revealed that R_e is related to the onset of turbulence. For R_e below about 2000, the flow is laminar. For R_e above about 3000, the flow is turbulent. For values of R_e between about 2000 and 3000, the flow is unstable – that is, it can be laminar, but small obstructions and surface roughness can make it turbulent, and it may oscillate randomly between being laminar and turbulent. The blood flow through most of the body is a quiet, laminar flow. The exception is in the aorta, where the speed of the blood flow rises above a critical value of 35 m/s and becomes turbulent.

An occlusion, or narrowing, is likely to cause turbulence because of the irregularity of the blockage, as well as the complexity of blood as a fluid. Turbulence in the circulatory system is noisy and can sometimes be detected with a stethoscope, such as when measuring diastolic pressure in the upper arm's partially collapsed brachial artery. These turbulent sounds, at the onset of blood flow when the cuff pressure becomes sufficiently small, are called *Korotkoff sounds.* Aneurysms, or ballooning of arteries, create significant turbulence and can sometimes be detected with a stethoscope. Heart murmurs, consistent with their name, are sounds produced by turbulent flow around damaged and insufficiently closed heart valves. Ultrasound can also be used to detect turbulence as a medical indicator in a process analogous to Doppler-shift radar used to detect storms.

2.6. Motion of an Object in a Viscous Fluid

A moving object in a viscous fluid is equivalent to a stationary object in a flowing fluid stream. The flow of the stationary fluid around a moving object may be laminar, turbulent, or a combination of the two. Laminar flow occurs mostly when the objects in the fluid are small, such as raindrops, pollen, and blood cells in plasma. One of the consequences of viscosity is a resistance force called viscous drag F_D that is exerted on a moving object. This force typically depends on the object's speed. Experiments have shown that for laminar flow ($R_e < 1$) the viscous drag is proportional to speed. For laminar flow around a sphere, F_D is proportional to fluid viscosity η, the object's characteristic size L (radius r for sphere), and its speed v. All of which makes sense – the more viscous the fluid and the larger the object, the more drag we expect. For the special case of a small sphere of radius r moving slowly in a fluid of viscosity η, the drag force F_D is given by Stoke's law

$$F_S = 6\pi r\eta v. \tag{2.16}$$

An interesting consequence of the increase in F_D with speed is that an object falling through a fluid will not continue to accelerate indefinitely (as it would if we neglect air resistance, for example). Instead, viscous drag increases, slowing acceleration until a critical speed called the *terminal speed*, is reached, and the acceleration of the object becomes zero. Once this happens, the object continues to fall at constant speed (the *terminal speed*). This is the case for particles of sand falling in the ocean, cells falling in a centrifuge, and sky divers falling through the air. Figure 2.7 shows some of the factors that affect terminal speed. There is a viscous drag on the object that depends on the viscosity of the fluid and the size of the object. But there is also a buoyant force that depends on the density of the object relative to the fluid. Terminal speed will be the greatest for low viscosity fluids and objects with high densities and small sizes.

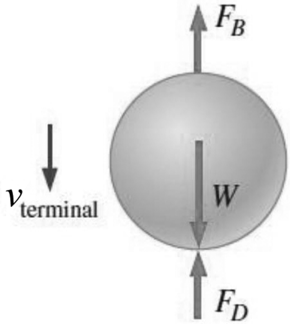

Figure 2.7 – There are three forces acting on an object falling through a viscous fluid: its weight W, the viscous drag F_D, and the buoyant force F_B

Knowledge of terminal speed is useful for estimating *sedimentation rates* of small particles. We know from watching mud settle out of dirty water that sedimentation is usually a slow process. Centrifuges are used to speed sedimentation by creating accelerated frames in which gravitational acceleration is replaced by centripetal acceleration, which can be much greater, increasing the terminal speed. From the free body diagram (see Figure 2.7), it is clear to see that $F_D = W - F_B$. The buoyancy force is simply the weight of displaced fluid. As you may recall from earlier works in science and math, the volume of a sphere (V_{sph}) is written as,

$$V_{shp} = \frac{4}{3}\pi r^3.$$

Combining this volume with the mass density of the fluid, ρ_{fl}, we can now write the buoyancy force as the product,

$$F_B = \frac{4}{3}\rho_{fl} g \pi r^3,$$

where g is the gravitational acceleration and r is the radius of the sphere. The drag force F_D is given by Stoke's law (2.16).

We can also write the weight of sphere in terms of its density ρ_{sph}:

$$W = m_{sph} g = \rho_{sph} V_{sph} g = \frac{4}{3}\rho_{sph} g \pi r^3.$$

Combining all of the previous relationships that describe the forces acting on the sphere in a fluid, we can write the following expression

$$6\pi r \eta v = \frac{4}{3}\rho_{sph} g \pi r^3 - \frac{4}{3}\rho_{fl} g \pi r^3.$$

Rearranging and regrouping the terms from the above equation, we arrive at the following relationship for the terminal speed of a slowly moving sphere in a viscous fluid

$$v = \frac{2r^2(\rho_{sph} - \rho_{fl})g}{9\eta}. \tag{2.17}$$

By measuring the terminal speed of a slowly moving sphere in a viscous fluid, one can find the viscosity of that fluid (at that temperature).

Lecture 3

Mechanical Properties of Solids

3.1. Elasticity: Stress and Strain

Deformation is a change in shape due to the application of a force. Even very small forces are known to cause some *deformation.* For small deformations, two important characteristics are observed. First, the object returns to its original shape when the force is removed – that is, the deformation is elastic for small deformations. Second, the size of the deformation is proportional to the force – that is, for small deformations, *Hooke's law* is obeyed. In equation form, Hooke's law is given by

$$F = k\Delta L, \tag{3.1}$$

where ΔL is the amount of deformation (the change in length, for example) produced by the force F, and k is a proportionality constant that depends on the shape and composition of the object and the direction of the force. Rearranging this to

$$\Delta L = \frac{F}{k} \tag{3.2}$$

makes it clear that the deformation is proportional to the applied force.

Figure 3.1 shows the Hooke's law relationship between the extension ΔL of a spring or of a human bone. For metals or springs, the straight line region in which Hooke's law pertains is much larger. Bones are brittle and the elastic region is small, and the abrupt fracture occurs. Eventually a large enough stress to the material will cause it to break or fracture.

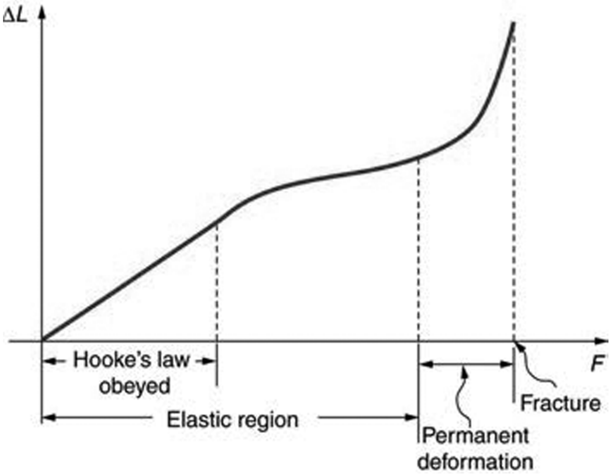

Figure 3.1 – A graph of deformation ΔL versus applied force F

The straight segment (see Figure 3.1) is the linear region where Hooke's law is obeyed. The slope of the straight region is $1/k$. For larger forces, the graph is curved but the deformation is still elastic: ΔL will return to zero if the force is removed. Still greater forces permanently deform the object until it finally fractures.

3.2. Tension and Compression. Elastic Modulus

A change in length ΔL is produced when a force is applied to a wire or rod parallel to its length L, either stretching it (a tension) or compressing it (see Figure 3.2).

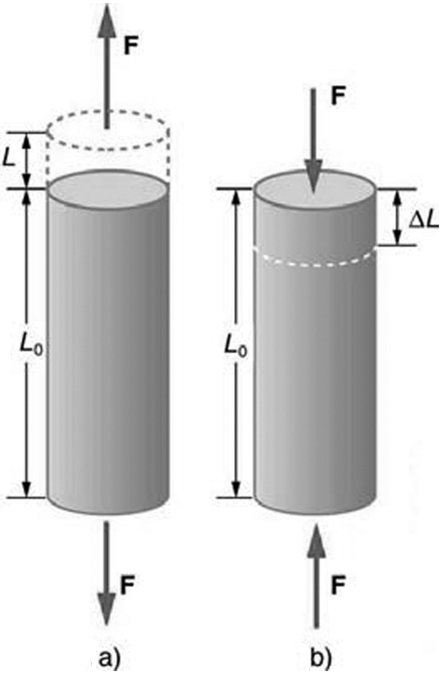

Figure 3.2 – a) Tension. The rod is stretched a length ΔL when a force is applied parallel to its length.
b) Compression

Experiments have shown that the change in length ΔL depends on only a few variables. As already noted, ΔL is proportional to the force F and depends on the substance from which the object is made. Additionally, the change in length is proportional to the original length L and inversely proportional to the cross-sectional area A of the wire or rod.

We can combine all these factors into one equation for ΔL:

$$\Delta L = \frac{1}{E}\frac{F}{A}L, \tag{3.3}$$

where E is a factor, called the elastic modulus or Young's modulus that depends on the substance. Table 3.1 lists values of E for several materials – those with a large E are said to have a large tensile strength because they deform less for a given tension or compression.

Bones, on the whole, do not fracture due to tension or compression. Rather they generally fracture due to sideways impact or bending, resulting in the bone shearing or snapping. The behavior of bones under tension and compression is important because it determines the load the bones can carry. Bones are classified as weight-bearing structures such as columns in buildings and trees. Weight-bearing structures have special features; columns in building have steel-reinforcing rods while trees and bones are fibrous. The bones in different parts of the body serve different structural functions and are prone to different stresses. Thus the bone in the top of the femur is arranged in thin sheets separated by marrow while in other places the bones can be cylindrical and filled with marrow or just solid. Overweight people have a tendency toward bone damage due to sustained compressions in bone joints and tendons.

Another biological example of Hooke's law occurs in tendons. Functionally, the tendon (the tissue connecting muscle to bone) must stretch easily at first when a force is applied, but offer a much greater restoring force for a greater strain. Some tendons have a high collagen content so there is relatively little strain, or length change; others, like support tendons (as in the leg) can

Table 3.1 – Elastic Moduli

Material	**Young's modulus** E**,** $10^9\ N/m^2$	**Shear modulus** S**,** $10^9\ N/m^2$	**Bulk modulus** B**,** $10^9\ N/m^2$
Aluminum	70	25	75
Bone (tension)	16	80	8
Bone (compression)	9		
Brass	90	35	75
Glass	70	20	30
Hair (human)	10		
Hardwood	15	10	
Iron	100	40	90
Steel	210	80	130
Lead	16	5	50
Nylon	5		
Silk	6		
Spider thread	3		
Tendon	1		
Acetone			0.7
Ethanol			0.9
Glycerin			4.5
Mercury			25
Water			2.2

change length up to 10%. Note that this stress-strain curve is nonlinear, since the slope of the line changes in different regions. In the first part of the stretch called the toe region, the fibers in the tendon begin to align in the direction of the stress-this is called uncrimping. In the linear region, the fibrils will be stretched, and in the failure region, individual fibers begin to break. A simple model of this relationship can be illustrated by springs in parallel: different springs are activated at different lengths of stretch. Ligaments (tissue connecting bone to bone) behave in a similar way.

Unlike bones and tendons, which need to be strong as well as elastic, the arteries and lungs need to be very stretchable. The elastic properties of the arteries are essential for blood flow. The pressure in the arteries increases and arterial walls stretch when the blood is pumped out of the heart. When the aortic valve shuts, the pressure in the arteries drops and the arterial walls relax to maintain the blood flow. When you feel your pulse, you are feeling exactly this – the elastic behavior of the arteries as the blood gushes through with each pump of the heart. If the arteries were rigid, you would not feel a pulse.

The heart is also an organ with special elastic properties. The lungs expand with muscular effort when we breathe in but relax freely and elastically when we breathe out. Our skins are particularly elastic, especially for the young. The elasticity of all organs reduces with age.

3.3. Stress and Strain

The equation for change in length is traditionally rearranged and written in the following form:

$$\frac{F}{A} = E\frac{\Delta L}{L}. \tag{3.4}$$

The ratio of force to area

$$\sigma = \frac{F}{A} \tag{3.5}$$

is defined as *stress* (measured in N/m^2), and the ratio of the change in length to length

$$\epsilon = \frac{\Delta L}{L} \tag{3.6}$$

is defined as strain (a unitless quantity).

In other words,

$$\sigma = E\epsilon. \tag{3.7}$$

3.4. Sideways Stress – Shear Stress

Figure 3.3 illustrates what is meant by a sideways stress or a *shearing force.* Here the deformation is called Δx and it is perpendicular to L, rather than parallel as with tension and compression. Shear deformation behaves similarly to tension and compression and can be described with similar equations. The expression for *shear deformation* is

$$\Delta x = \frac{1}{S}\frac{F}{A}L, \tag{3.8}$$

where S is the shear modulus (see Table 3.1) and F is the force applied perpendicular to L and parallel to the cross-sectional area A.

In other words,

$$\tau = S\phi, \tag{3.9}$$

where $\tau = F/A$ *shear stress* (measured in N/m^2), and the ratio of shear deformation to length

$$\phi = \frac{\Delta X}{L} \tag{3.10}$$

is defined as shear strain (a unitless quantity).

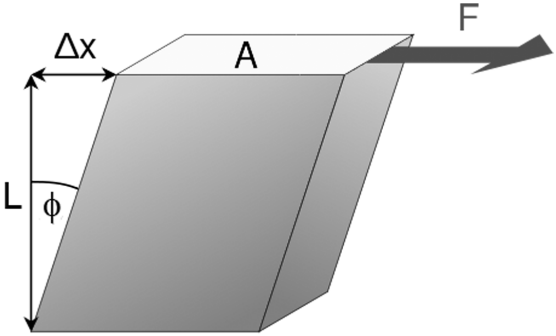

Figure 3.3 – Shearing forces are applied perpendicular to the length L and parallel to the area A, producing a deformation Δx

Examination of the shear moduli in Table 3.1 reveals some telling patterns. For example, shear moduli are less than Young's moduli for most materials. Bone is a remarkable exception. Its shear modulus is not only greater than its Young's modulus, but it is as large as that of steel. This is one reason that bones can be long and relatively thin. Bones can support loads comparable to that of concrete and steel. Most bone fractures are not caused by compression but by excessive twisting and bending. The spinal column (consisting of 26 vertebral segments separated by discs) provides the main support for the head and upper part of the body. The spinal column has normal curvature for stability, but this curvature can be increased, leading to increased shearing forces on the lower vertebrae. Discs are better at withstanding compressional forces than shear forces. Because the spine is not vertical, the weight of the upper body exerts some of both. Pregnant women and people that are overweight (with large abdomens) need to move their shoulders back to maintain balance, thereby increasing the curvature in their spine and so increasing the shear component of the stress. An increased angle due to more curvature increases the shear forces along the plane. These higher shear forces increase the risk of back injury through ruptured discs. The lumbosacral disc (the wedge shaped disc below the last vertebrae) is particularly at risk because of its location.

3.5. Volume's Changes. Bulk Modulus

An object will be compressed in all directions if inward forces are applied evenly on all its surfaces as in Figure 3.4. It is relatively easy to compress gases and extremely difficult to compress liquids and solids. For example, air in a wine bottle is compressed when it is corked. But if you try corking a brim-full bottle, you cannot compress the wine – some must be removed if the cork is to be inserted. The reason for these different compressibilities is that atoms and molecules are separated by large empty spaces in gases but packed close together in liquids and solids. To compress a gas, you must force its atoms and molecules to drive closer together. To compress liquids and solids, you must actually compress their atoms and molecules, and very strong electromagnetic forces in them oppose this compression.

We can describe the compression or volume deformation of an object with an equation. First, we note that a force "applied evenly" is defined to have the same stress or ratio of force to area F/A on all surfaces. The deformation produced is a change in volume ΔV, which is found to behave very similarly to the shear, tension, and compression previously discussed. (This is not surprising since a compression of the entire object is equivalent to

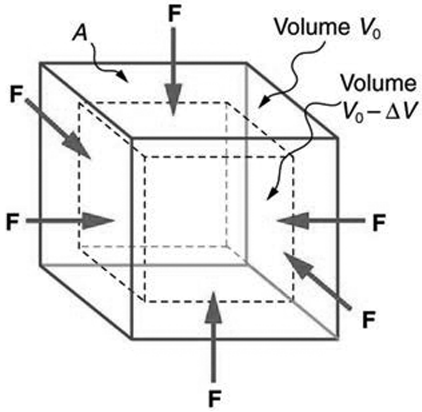

Figure 3.4 – An inward force on all surfaces compresses this cube. Its change in volume is proportional to the force per unit area and its original volume, and is related to the compressibility of the substance

compressing each of its three dimensions.) The relationship of the change in volume to other physical quantities is given by

$$\Delta V = \frac{1}{B}\frac{F}{A}V, \tag{3.11}$$

where B is the bulk modulus (see Table 3.1), V is the original volume, and F/A is the force per unit area applied uniformly inward on all surfaces. Note that no bulk moduli are given for gases.

3.6. Viscoelasticity

Mechanical Properties of the Tissues

The soft biological tissues (skin, tendon, ligament) play an important role in the mechanical integrity of the body. Indeed, these tissues have the following functions: to protect the body for the skin, to transfer loads between bones for the ligaments, or between muscles and bones for the tendons. The soft biological tissues are mainly made of collagen and elastin proteins, which bring special mechanical properties. The tissues can be stretched 15% without damage. They also have an important viscous component in their behaviours.

A primary group of tissue which binds, supports and protects our human body and structures such as organs is soft connective tissue. In contrary to other tissues, it is a wide-ranging biological material in which the cells are separated by extracellular material. Connective tissues may be distinguished from hard (mineralized) tissues, such as bones, for their high flexibility and their soft mechanical properties.

Examples for soft tissues are tendons, ligaments, blood vessels, skins or articular cartilages among many others. Tendons are muscle-to-bone linkages to stabilize the bony skeleton (or to produce motion), while ligaments are bone-to-bone linkages to restrict relative motion. Blood vessels are prominent organs composed of soft tissues which have to distend in response to pulse waves. The skin is the largest single organ (16% of the human adult weight). It supports internal organs and protects our body. Articular cartilages form the surface of body joints (which is a layer of connective tissue with a thickness of 1-5 mm) and distribute loads across joints and minimize contact stresses and friction.

Soft connective tissues of our body are complex fiber-reinforced composite structures. Their mechanical behavior is strongly influenced by the concentration and structural arrangement of constituents such as *collagen* and *elastin*, the hydrated matrix of *proteoglycans*, and the topographical site and respective function in the organism.

Collagen: Collagen is a protein which is a major constituent of the extracellular matrix of connective tissue. It is the main load carrying element in a wide variety of soft tissues and is very important to human physiology (for example, the collagen content of (human) achilles tendon is about 20 times that of elastin).

Collagen is a macromolecule with length of about 280 nm. Collagen molecules are linked to each other by covalent bonds building collagen fibrils. Depending on the primary function and the requirement of strength of the tissue the diameter of collagen fibrils varies (the order of magnitude is 1.5 nm). In the structure of tendons and ligaments, for example, collagen appears as parallel oriented fibers, while many other tissues have an intricate disordered network of collagen fibers embedded in a gelatinous matrix of proteoglycans.

Elastin: Elastin, like collagen, is a protein which is a major constituent of the extracellular matrix of connective tissue. It is present as thin strands in soft tissues such as skin, lung, ligamenta flava of the spine and ligamentum nuchae (the elastin content of the latter is about 5 times that of collagen).

The long flexible elastin molecules build up a three-dimensional (rubber-like) network, which may be stretched to about 2.5 of the initial length of the unloaded configuration. In contrast to collagen fibers, this network does not exhibit a pronounced hierarchical organization. As for collagen, 33% of the total amino acids of elastin consists of glycine. However, the proline and hydroxyproline contents are much lower than in collagen molecules.

Viscoelastic Models

Soft tissues and cells exhibit several anelastic properties: *hysteresis* during loading and unloading; *stress relaxation* at constant strain; *creep* at constant stress; *strain-rate dependence.* In general, stress in soft tissues depends on strain and the history of strain. These properties can be modeled by the theory of viscoelasticity.

Hysteresis: When a body is subjected to cyclic loading, load-displacement (or stress-strain) behavior for increasing loads is different than behavior for decreasing loads. The area between the curves represents energy loss (dissipation).

Stress relaxation: When a body is deformed (or strained) and that deformation (or strain) is held constant, stresses in the body reduce with time. The stress relaxation test, easily conducted on displacement-controlled machines, consists of monitoring the time-dependent stress resulting from a steady strain.

Creep: When a body is loaded (or stressed) and the stress is held constant, the body continues to deform (or strain) with time. The creep test consists of measuring the time dependent strain $\epsilon(t)$ resulting from the application of a steady uniaxial stress σ.

For viscoelastic materials, the relationship between stress and strain can be expressed as

$$\sigma = \sigma(\epsilon, \dot{\epsilon}).$$

We used short-hand notation of time derivative called the "over-dot"($\dot{x} = \frac{dx}{dt}$).

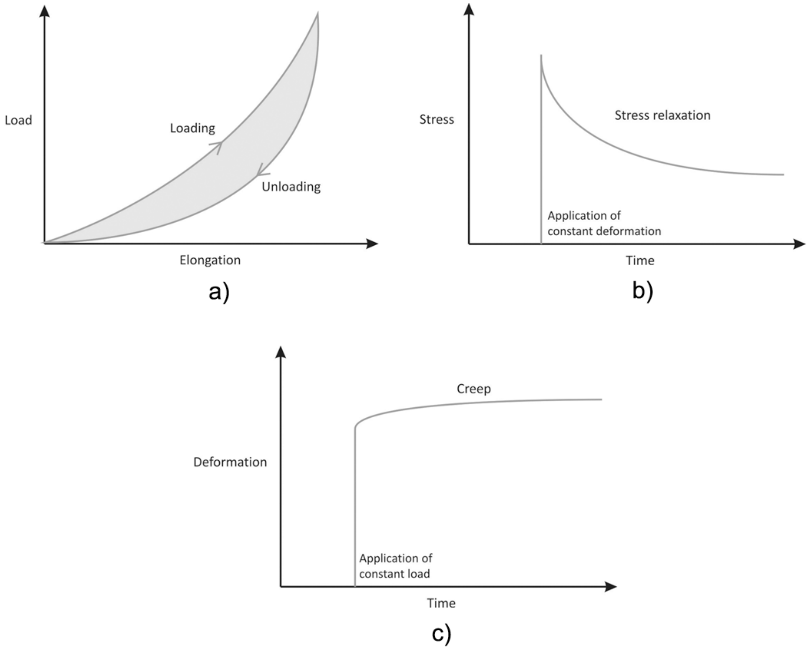

Figure 3.5 – a) Hysteresis, b) Stress relaxation, c) Creep

Behavior exhibited by a material (or tissue) that has both viscous and elastic elements in its response to a deformation (or strain) or load (or stress) represented by a linear spring (see Figure 3.6) with Young's modulus E theoretically produces a deformation proportional to load $\sigma = E\epsilon$; a dashpot with coefficient of viscosity η produces a velocity proportional to load $\sigma = \eta\dot{\epsilon}$ (Newtonian fluid constitutive law).

Simple Linear Viscoelastic Models

The viscoelastic models are all composed of combinations of linear springs and dashpots.

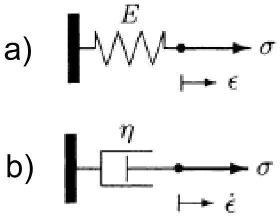

Figure 3.6 – a) Elastic element (spring), b) Viscous element (dashpot)

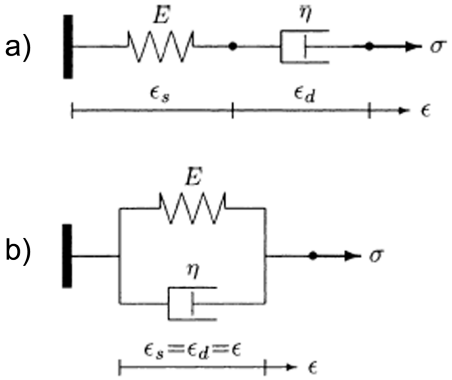

Figure 3.7 – a) Maxwell model, b) Kelvin-Voigt model

Maxwell Model: It is represented by a purely viscous damper and a purely elastic spring connected in series.

The model can be represented by the following differential equation:

$$\dot{\epsilon} = \frac{\sigma}{\eta} + \frac{\dot{\sigma}}{E}. \tag{3.12}$$

Maxwell model predicts a stress that decays exponentially with time to zero with permanent deformation. Model doesn't

Table 3.2 – Maxwell and Kelvin-Voigt model comparison

Maxwell model	**Kelvin-Voigt model**
Creep functions ($\sigma = \sigma_0 = const$)	
$\epsilon(t) = \frac{\sigma_0}{\eta} t$	$\epsilon(t) = \frac{\sigma_0}{E}\left(1 - e^{-\frac{E}{\eta}t}\right)$
Relaxation functions ($\epsilon = \epsilon_0 = const$)	
$\sigma(t) = \sigma_0^{-\frac{E}{\eta}t}$	$\sigma(t) = E\epsilon_0$

accurately predict creep (constant stress). It predicts that strain will increase linearly with time. Actually, strain rate decreases with time.

Kelvin-Voigt Model Represented by a Newtonian damper and Hookean elastic spring in parallel.

The model can be expressed as a linear first order differential equation

$$\sigma = E\epsilon + \eta\dot{\epsilon}. \tag{3.13}$$

The model represents a solid undergoing reversible, viscoelastic strain; a solid that is very stiff but will creep (e.g., crystals, glass, apparent behavior of cartilage). At constant stress (creep), it predicts strain to tend to σ/E as time continues to infinity. The model is not accurate for relaxation in a material (tissue).

Lecture 4

Mechanical Oscillations and Waves

4.1. Simple Harmonic Motion

Oscillatory motion is everywhere in nature. Any object which has both inertia and a restoring force will oscillate around an equilibrium position if displaced from that equilibrium. The time T to complete one oscillation remains constant and is called the *period*. Its units are usually seconds, but may be any convenient unit of time. Frequency f is defined to be the number of events per unit time. For periodic motion, frequency is the number of oscillations per unit time. The relationship between frequency and period is

$$f = \frac{1}{T}. \tag{4.1}$$

The SI unit for frequency is the *cycle per second*, which is defined to be a hertz (Hz): $1\,Hz = s^{-1}$. A cycle is one complete oscillation.

The oscillations of a system in which the net force can be described by Hooke's law are of special importance, because they are very common. They are also the simplest oscillatory systems. *Simple Harmonic Motion* (SHM) is the name given to oscillatory motion for a system where the net force can be described by Hooke's law, and such system is called a *simple harmonic oscillator* (SHO). If the net force can be described by Hooke's law and there is no damping (by friction or other non-conservative forces), then a simple harmonic oscillator will oscillate with equal displacement on either side of the equilibrium position, as shown for an object on a spring in Figure 4.1. The deformation of the spring creates a force in the opposite direction, known as a *restoring force.*

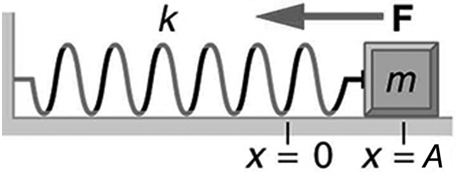

Figure 4.1 – An object attached to a spring sliding on a frictionless surface is an uncomplicated simple harmonic oscillator.

According to Hooke's law, the force required to stretch or compress the spring is proportional to the amount the spring stretches or compresses: $F = -kx$. The force constant k is related to the rigidity (or stiffness) of a system – the larger the force constant, the greater the restoring force, and the stiffer the system. The units of k are newtons per meter (N/m).

If a mass m is attached to the end of the spring, displaced from equilibrium and released, the only unbalanced force acting on the mass is that exerted by the spring. Newton's second law can then be expressed as

$$ma = -kx.$$

This can be expressed in differential form:

$$m\frac{d^2x}{dt^2} = -kx.$$

The solution to this second order differential equation is either a sine or a cosine function with time as the variable. In general, the solution can be written as a sine or cosine function with an arbitrary *phase* (the angle ϕ_0) in the argument.

$$x(t) = A\cos(\omega_0 t + \phi_0), \tag{4.2}$$

where $\omega_0 = \sqrt{k/m}$ is the angular frequency of SHM. The angular frequency relates to the frequency of the oscillation by $\omega_0 = 2\pi f$. Notice that the role of the phase ϕ is to identify where the object is with respect to the equilibrium position at time $t = 0$.

The maximum displacement from equilibrium is called the *amplitude* A. The units for amplitude and displacement are the same but depend on the type of oscillation. For the object on the spring, the units of amplitude and displacement are meters; whereas for sound oscillations, they have units of pressure (and other types of oscillations have yet other units). Because amplitude is the maximum displacement, it is related to the energy in the oscillation.

If the displacement is described by (4.2), then the velocity and acceleration, respectively, will be given by

$$v(t) = -\omega_0 A\sin(\omega_0 t + \phi_0) \quad \text{and} \quad a(t) = -\omega_0^2 A\cos(\omega_0 t + \phi_0), \tag{4.3}$$

where $\omega_0 A = v_{max}$ is the amplitude of velocity, $\omega_0^2 A = a_{max}$ is the amplitude of acceleration.

Figure 4.2 shows the simple harmonic motion of an object on a spring and presents graphs of $x(t)$, $v(t)$, and $a(t)$ versus time.

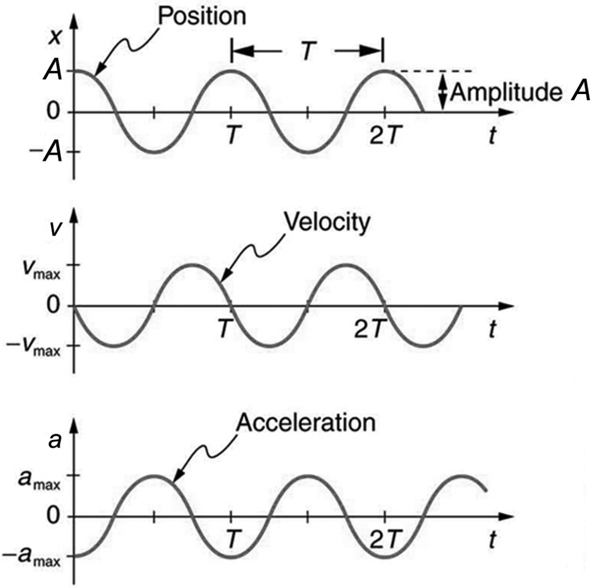

Figure 4.2 – Graphs of $x(t)$, $v(t)$, and $a(t)$ versus t for the motion of an object on a spring

4.2. Energy of a SHO

The energy associated with a harmonic oscillator is just the sum of the kinetic and potential energies. For a mass on a spring, that is just

$$E_{total} = E_K + E_P.$$

For the mass-spring system, the kinetic energy is given by

$$E_K = \frac{1}{2}mv^2 = \frac{1}{2}\omega_0^2 A^2 \sin^2(\omega_0 t + \phi_0).$$

The elastic potential energy stored in any deformed system that obeys Hooke's law and has a displacement x from equilibrium and a force constant k is given by

$$E_P = \frac{1}{2}kx^2 = \frac{1}{2}kA^2\cos^2(\omega_0 t + \phi_0).$$

Substituting $k = m\omega_0^2$ reduces the above to

$$E_P = \frac{1}{2}m\omega_0^2A^2\cos^2(\omega_0 t + \phi_0).$$

The total energy of the simple harmonic oscillator is

$$E_{total} = \frac{1}{2}m\omega_0^2A^2\left(\sin^2(\omega_0 t + \phi_0) + \cos^2(\omega_0 t + \phi_0)\right),$$

and the trig identity $\sin^2(\alpha) + \cos^2(\alpha) = 1$ gives

$$E_{total} = \frac{1}{2}m\omega_0^2A^2, \tag{4.4}$$

showing that the energy of the simple harmonic oscillator (as typified by a mass on a spring) is constant and is equal to the potential energy of the spring when it is maximally extended (at which time the mass is motionless).

4.3. Damped Harmonic Motion

Consider a mass-spring system which has some damping mechanism which depends on the speed of the oscillator: for example, a system which moves in a viscous fluid would provide such a mechanism.

The resistive force depends on the speed of the motion and can be expressed as

$$F_R = -rv,$$

and it is in addition to the restoring force of the spring. So Newton's second law for such an oscillator can be written:

$$m\frac{d^2x}{dt^2} = -kx - rv.$$

Solving for the differential equation yields

$$x(t) = A(t)\cos(\omega t + \phi_0), \tag{4.5}$$

where $A(t) = A_0 e^{-\beta t}$ is the amplitude of oscillation, $\beta = r/(2m)$ is damping coefficient, and ω is angular frequency of damped oscillation which is given by

$$\omega = \sqrt{\omega_0^2 - \beta^2}. \tag{4.6}$$

The primary effect of the damping force on the motion of the oscillator – assuming the damping is not so great as to prevent it from oscillating at all (a situation called "over-damping") – is to cause the amplitude of the oscillation to diminish in time (see Figure 4.3). The time dependence of the amplitude is given by the decreasing exponential function.

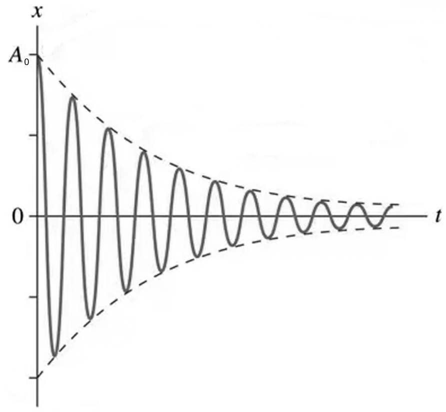

Figure 4.3 – Graph of displacement versus time for a harmonic oscillator with a small amount of damping

4.4. Forced Oscillations and Resonance

In this section, we shall briefly explore applying a periodic driving force acting on a simple harmonic oscillator. The driving force puts energy into the system at a certain frequency, not necessarily the same as the natural frequency of the system. The *natural frequency* is the frequency at which a system would oscillate if there were no driving or damping force. The phenomenon of driving a system with a frequency equal to its natural frequency is called *resonance.* The resonant angular frequency is determined by formula

$$\omega_{res} = \sqrt{\omega_0^2 - 2\beta^2}. \tag{4.7}$$

A system being driven at its natural frequency is said to *resonate.* As the driving frequency gets progressively higher than the resonant or natural frequency, the amplitude of the oscillations becomes smaller, until the oscillations nearly disappear as shown in Figure 4.4.

These features of driven harmonic oscillators may be applied to a huge variety of systems. When you tune a radio, for example, you are adjusting its resonant frequency so that it only oscillates to the desired station's broadcast (driving) frequency. The more selective the radio is in discriminating between stations, the smaller its damping. Magnetic resonance imaging (MRI) is a widely used medical diagnostic tool in which atomic nuclei (mostly hydrogen nuclei) are made to resonate by incoming radio waves (on the order of 100 MHz). A child on a swing is driven by a parent at the natural frequency of the swing to achieve maximum amplitude.

In our body, the chest cavity is a clear example of a system at resonance. The diaphragm and chest wall drive the oscillations of the chest cavity which result in the lungs inflating and deflating. The system is critically damped and the muscular diaphragm

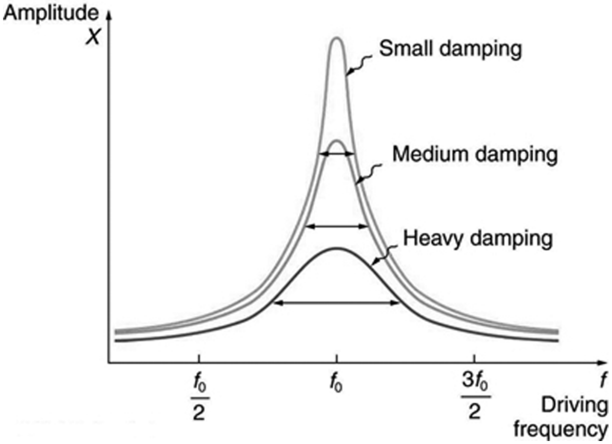

Figure 4.4 – Amplitude of a harmonic oscillator as a function of the frequency of the driving force

oscillates at the resonant value for the system, making it highly efficient.

4.5. Waves

A *wave* is a disturbance that propagates or moves from the place it was created. We define *wave velocity* v to be the speed at which the disturbance moves. Wave velocity is sometimes also called the *propagation velocity* or *propagation speed*, because the disturbance propagates from one location to another.

A wave also has a length associated with it, called its *wavelength* λ, the distance between adjacent identical parts of a wave. (λ is the distance parallel to the direction of propagation.) The speed of propagation v is the distance the wave travels in a given time, which is one wavelength in the time of one period. In equation form, that is,

$$v = \frac{\lambda}{T} = \lambda f. \tag{4.8}$$

This fundamental relationship holds for all types of waves. For water waves, v is the speed of a surface wave; for sound, v is the speed of sound; and for visible light, v is the speed of light, for example.

Transverse and Longitudinal Waves

A simple wave consists of a periodic disturbance that propagates from one place to another. The wave in Figure 4.5 propagates in the horizontal direction while the surface is disturbed in the vertical direction. Such a wave is called a *transverse wave* or *shear wave*; in such a wave, the disturbance is perpendicular to the direction of propagation. In contrast, in a *longitudinal wave* or *compressional wave*, the disturbance is parallel to the direction of propagation. Figure 4.5 shows an example of a longitudinal wave. The size of the disturbance is its amplitude A and is completely independent of the speed of propagation v.

Waves may be transverse, longitudinal, or a combination of the two. (Water waves are actually a combination of transverse and longitudinal waves. The waves on the strings of musical instruments are transverse – so are electromagnetic waves, such as visible light. Sound waves in air and water are longitudinal waves. Their disturbances are periodic variations in pressure that are transmitted in fluids.

Fluids do not have appreciable shear strength, and thus the sound waves in them must be longitudinal or compressional. Sound waves in solids can be both longitudinal and transverse.

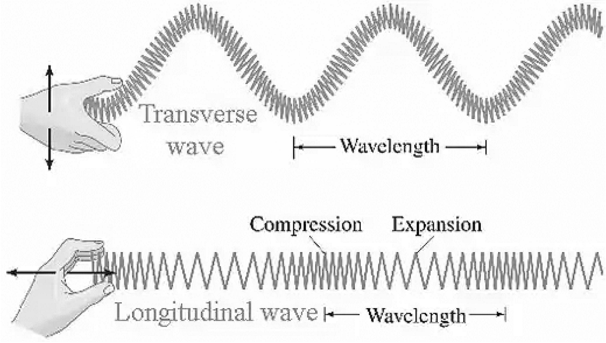

Figure 4.5 – Example of a transverse and longitudinal waves.

Moving Waves

There is no obvious connection between the motion of a wave and the motion of particles in the medium through which the wave travels. General equation of traveling wave in a medium is given by

$$y(x,t) = A\cos(\omega t - kx + \phi_0), \tag{4.9}$$

where y is a particle of a medium displacement – a distance that the medium particle is moved from its equilibrium position at any time t, A is the amplitude of wave – maximum particle displacement from its equilibrium position, ω is angular frequency of oscillation, $k = \omega/v$ is the wave number, x is the stationary coordinate and ϕ_0 is an arbitrary angle (called a phase constant). Note that the equation (4.9) is written for the wave which is traveling in positive x-direction.

The Standing Wave in a String

When a transverse wave is generated in a taut string of length l, two travelling waves are formed; these should have the same amplitudes, only they travel in opposite directions. The displacements of overlapping waves simply add (or subtract) together, and so the wave in the string has the form

$$y_s(x,t) = A\cos(\omega t - kx) + A\cos(\omega t - kx) = 2A\cos(\omega t)\sin(kx). \tag{4.10}$$

This last equation does not describe traveling waves, it describes waves that are fixed on the x axis by $\sin(kx)$ while their transverse (sideways) displacement changes with time according to $2A\cos(\omega t)$.

The fixed parts of the string must correspond with a fixed zero in our functions; these zero points are called *nodes.* The word *antinode* is used to denote the location of maximum amplitude in standing waves.

We can see (Figure 4.6) that the distances between nodes, say Δx, are found from $k\Delta x = \pi$, 2π, 3π or more generally $n\pi$, where n is a positive integer. This relation gives us the standing wavelengths in the string as

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

so that

$$\lambda_n = \frac{2l}{n}.$$

We have a discrete number of choices for the wavelength of a transverse wave in a string of length l, each particular value of n defines a *mode.* For a *fundamental* mode or *first harmonic* $n = 1$. For the *second harmonic* mode $n = 2$, for the *third harmonic* mode $n = 3$, etc. As the speed of a wave v in a string depends on the tension in the string, we can also calculate the frequency of the modes (or harmonics):

$$f_n = \frac{v}{2l} n. \tag{4.11}$$

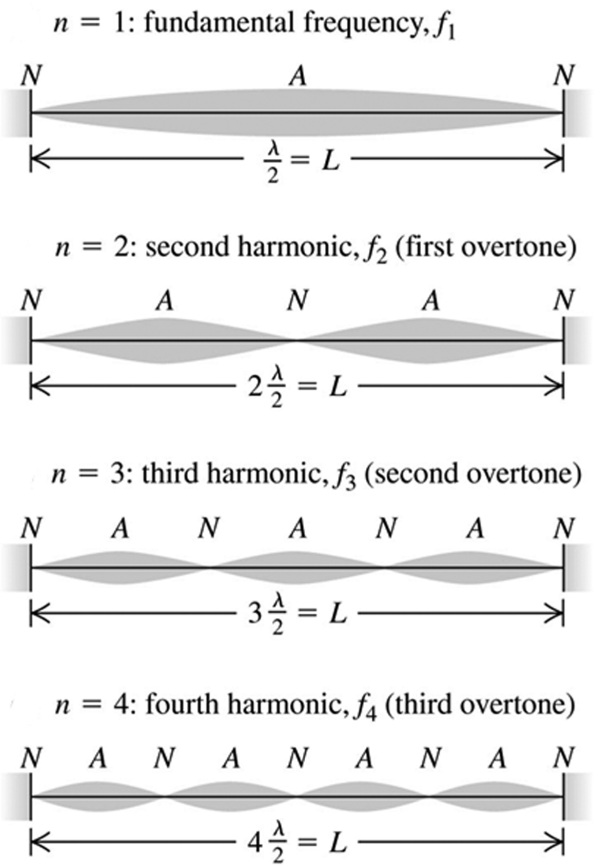

Figure 4.6 – The standing wave in the string

Superposition and Interference

When two or more waves arrive at the same point, they superimpose themselves on one another. More specifically, the disturbances of waves are superimposed when they come together: the phenomenon is called *superposition*. Each disturbance corresponds to a force, and forces add. If the disturbances are along the same line, then the resulting wave is a simple addition of the disturbances of the individual waves – that is, their amplitudes add.

Figure 4.7 shows two identical waves that arrive at the same point exactly in phase. The crests of the two waves are precisely aligned, as are the troughs. This superposition produces pure *constructive interference.* Because the disturbances add, pure constructive interference produces a wave that has twice the amplitude of the individual waves, but has the same wavelength.

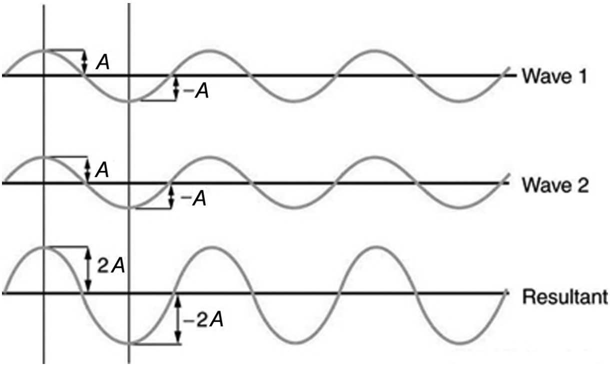

Figure 4.7 – Pure constructive interference of two identical waves produces one with twice the amplitude, but the same wavelength

Figure 4.8 shows two identical waves that arrive exactly out of phase – that is, precisely aligned crest to trough – producing pure *destructive interference.* Because the disturbances are in the opposite direction for this superposition, the resulting amplitude is zero for pure destructive interference – the waves completely cancel each other out.

The superposition of most waves produces a combination of constructive and destructive interference and can vary from place to place and time to time.

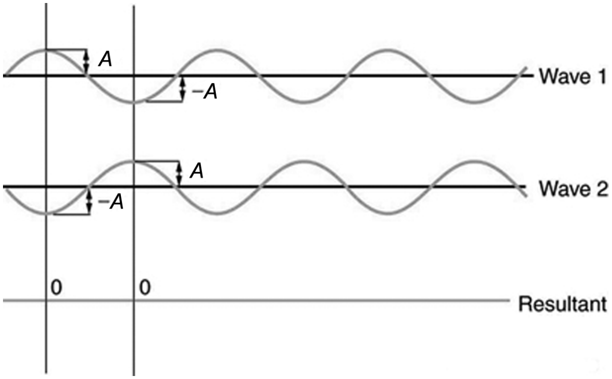

Figure 4.8 – Pure destructive interference of two identical waves produces zero amplitude, or complete cancellation

Beats

Striking two adjacent keys on a piano produces a warbling combination usually considered to be unpleasant. The superposition of two waves of similar but not identical frequencies is the culprit. Another example is often noticeable in jet aircraft, particularly the two-engine variety, while taxiing. The combined sound of the engines goes up and down in loudness. This varying loudness happens because the sound waves have similar but not identical frequencies. The discordant warbling of the piano and the fluctuating loudness of the jet engine noise are both due to alternately constructive and destructive interference as the two waves go in and out of phase. Figure 4.9 illustrates this graphically.

The wave resulting from the superposition of two similar-frequency waves has a frequency that is the average of the two. This wave fluctuates in amplitude, or beats, with a frequency called the *beat frequency*. Adding two waves that have different

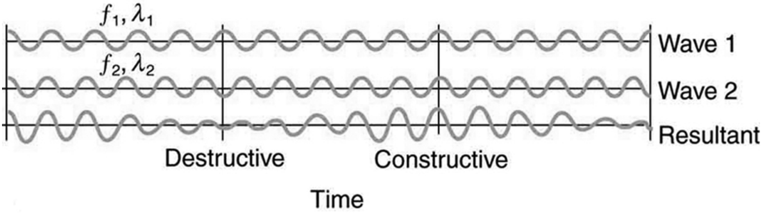

Figure 4.9 – Beats are produced by the superposition of two waves of slightly different frequencies but identical amplitudes. The waves alternate in time between constructive interference and destructive interference, giving the resulting wave a time-varying amplitude.

frequencies but identical amplitudes produces a resultant

$$y = y_1 + y_2 = A\cos(2\pi f_1 t) + A\cos(2\pi f_2 t).$$

Using a trigonometric identity, it can be shown that

$$y = 2A\cos(\pi f_B t)\cos(2\pi f_{ave} t),$$

where

$$f_B = |f_1 - f_2| \tag{4.12}$$

is the beat frequency, and f_{ave} is the average of f_1 and f_2. These results mean that the resultant wave has twice the amplitude and the average frequency of the two superimposed waves, but it also fluctuates in overall amplitude at the beat frequency f_B. The first cosine term in the expression effectively causes the amplitude to go up and down. The second cosine term is the wave with frequency $f_a ve$. This result is valid for all types of waves. However, if it is a sound wave, providing the two frequencies are similar, then what we hear is an average frequency that gets louder and softer (or warbles) at the beat frequency.

Energy in Waves: Intensity

All waves carry energy. The amount of energy in a wave is related to its amplitude. A wave's energy is directly proportional to its amplitude squared. The energy effects of a wave depend on time as well as amplitude. For example, the longer deep-heat ultrasound is applied, the more energy it transfers. Waves can also be concentrated or spread out. Sunlight, for example, can be focused to burn wood. In both cases, changing the area the wave covers has important effects. All these pertinent factors are included in the definition of *intensity* I as the power per unit area:

$$I = \frac{P}{A}, \tag{4.13}$$

where P is the power carried by the wave through area A. The definition of intensity is valid for any energy in transit, including that carried by waves. The SI unit for intensity is watts per square meter (W/m^2). For example, infrared and visible energy from the Sun impinges upon the Earth with the intensity of 1300 W/m^2 just above the atmosphere.

Lecture 5

Sound. Physics of Hearing

If a tree falls in the forest and no one is there to hear it, does it make a sound? The answer to this old philosophical question depends on how you define sound. If sound only exists when someone is around to perceive it, then there was no sound. However, if we define sound in terms of physics; that is, a disturbance of the atoms in matter transmitted from its origin outward (in other words, a wave), then there was a sound, even if nobody was around to hear it.

Such a wave is the physical phenomenon we call sound. Its perception is hearing. Both the physical phenomenon and its perception are interesting and will be considered in this lecture. We shall explore both sound and hearing; they are related, but are not the same thing. We will also explore the many practical uses of sound waves, such as in medical imaging.

5.1. Sound

Sound can be used as a familiar illustration of waves. Because hearing is one of our most important senses, it is interesting to see how the physical properties of sound correspond to our perceptions of it. *Hearing* is the perception of sound, just as vision is the perception of visible light. But sound has important applications beyond hearing. Ultrasound, for example, is not heard but can be employed to form medical images and is also used in treatment.

The physical phenomenon of *sound* is defined to be a disturbance of matter that is transmitted from its source outward. Sound is a wave. On the atomic scale, it is a disturbance of atoms that is far more ordered than their thermal motions. In many instances, sound is a periodic wave, and the atoms undergo simple harmonic motion.

Sound waves in air and most fluids are longitudinal, because fluids have almost no shear strength. In solids, sound waves can be both transverse and longitudinal.

The amplitude of a sound wave decreases with distance from its source, because the energy of the wave is spread over a larger and larger area. But it is also absorbed by objects, such as the eardrum, and converted to thermal energy by the viscosity of air. Wavelength, frequency, amplitude, and speed of propagation are important for sound, as they are for all waves.

5.2. Speed of Sound

Sound, like all waves, travels at a certain speed and has the properties of frequency and wavelength. You can observe direct evidence of the speed of sound while watching a fireworks display. The flash of an explosion is seen well before its sound is heard, implying both that sound travels at a finite speed and that it is much slower than light. You can also directly sense the frequency of a sound. Perception of frequency is called *pitch*. The wavelength of sound is not directly sensed, but indirect evidence is found in the correlation of the size of musical instruments with their pitch. Small instruments, such as a piccolo, typically make high-pitch sounds, while large instruments, such as a tuba, typically make low-pitch sounds. High pitch means small wavelength, and the size of a musical instrument is directly related to the wavelengths of sound it produces. So a small instrument creates short-wavelength sounds. Similar arguments hold that a large instrument creates long-wavelength sounds. The relationship of the speed of sound, its frequency, and wavelength is the same as for all waves:

$$v = \lambda f, \tag{5.1}$$

where v is the speed of sound, f is its frequency, and λ is its wavelength. The wavelength of a sound is the distance between adjacent identical parts of a wave.

The speed of sound depends on the characteristics of the medium that sound travels through: elasticity, density and temperature.

The speed of sound in a medium is determined by a combination of the medium's rigidity (or compressibility in gases) and its density. The more rigid (or less compressible) the medium, the faster the speed of sound. The greater the density of a medium, the slower the speed of sound. The speed of sound in air is low, because air is compressible. Because liquids and solids are relatively rigid and very difficult to compress, the speed of sound in such media is generally greater than in gases. The speed of sound in a medium is determined by formula:

$$v = \sqrt{\frac{B}{\rho}}, \tag{5.2}$$

where B is bulk modulus of a medium that basically tells you how hard it is to compress it, ρ is density of the medium.

The bulk modulus for air is tiny compared to that of water, since air is easily compressed and water nearly incompressible. So, even though water is much denser than air, water is so much harder to compress that sound travels over 4 times faster in water. Steel is almost 8 times denser than water, but it's over 70 times harder to compress. Consequently, sound waves propagate through steel about 3 times faster than in water. Table 5.1 makes it apparent that the speed of sound varies greatly in different media.

Table 5.1 – Speed of Sound in Various Media

Medium	v, m/s
Gases at $0°C$	
Air	331
Carbon dioxide	259
Oxygen	316
Helium	965
Hydrogen	1290
Liquids at $20°C$	
Ethanol	1160
Mercury	1450
Water	1480
Sea water	1540
Human tissue	1540
Solids (longitudinal)	
Vulcanized rubber	54
Polyethylene	920
Marble	3810
Glass, Pyrex	5640
Lead	1960
Aluminum	5120
Steel	5960

The speed of sound is affected by temperature in a given medium. For air at sea level, the speed of sound is given by

$$v = (331\,m/s)\sqrt{\frac{T}{273\,K}}, \tag{5.3}$$

where the temperature (denoted as T) is in units of kelvin.

The bulk modulus for air is tiny compared to that of water, since air is easily compressed and water nearly incompressible. So, even though water is much denser than air, water is so much harder to compress that sound travels over 4 times faster in water. Steel is almost 8 times denser than water, but it's over 70 times harder to compress. Consequently, sound waves propagate through steel about 3 times faster than in water.

The speed of sound can change when sound travels from one medium to another. However, the frequency usually remains the same because it is like a driven oscillation and has the frequency of the original source. If v changes and f remains the same, then the wavelength λ must change. That is, because $v = f\lambda$, the higher the speed of a sound, the greater its wavelength for a given frequency.

5.3. Sound Intensity and Sound Level

The relevant physical quantity is *sound intensity*, a concept that is valid for all sounds whether or not they are in the audible range. Intensity is defined to be the power per unit area carried by a wave. Power is the rate at which energy is transferred by the wave. In equation form, intensity I is

$$I = \frac{P}{A}, \tag{5.4}$$

where P is the power through an area A. The SI unit for I is W/m^2. The intensity of a sound wave is related to its amplitude squared by the following relationship:

$$I = \frac{\Delta p^2}{2\rho v}. \tag{5.5}$$

Here Δp is the pressure variation or pressure amplitude (half the difference between the maximum and minimum pressure in the sound wave) in units of pascals Pa or N/m^2. (We are using a lower case p for pressure to distinguish it from power, denoted by P above.)

Sound intensity levels are quoted in decibels (dB) much more often than sound intensities in watts per meter squared. Decibels are the unit of choice in the scientific literature as well as in the popular media. The reasons for this choice of units are related to how we perceive sounds. How our ears perceive sound can be more accurately described by the logarithm of the intensity rather than directly to the intensity. The sound intensity level β in decibels of a sound having an intensity I in watts per meter squared is defined to be

$$\beta(dB) = 10 \log \left(\frac{I}{I_0} \right), \tag{5.6}$$

where $I_0 = 10^{-12}\, W/m^2$ is a reference intensity. In particular, I_0 is the lowest or threshold intensity of sound a person with normal hearing can perceive at a frequency of 1000 Hz. Sound intensity level is not the same as intensity. The decibel level of a sound having the threshold intensity of $10^{-12}\, W/m^2$ is $\beta = 0\, dB$, because $\log_1 01 = 0$. That is, the threshold of hearing is 0 decibels. Table 5.2 gives levels in decibels and intensities in watts per meter squared for some familiar sounds.

One of the more striking things about the intensities in Table 5.2 is that the intensity in watts per meter squared is quite small for most sounds. The ear is sensitive to as little as a trillionth of a watt per meter squared. This is even more impressive when you realize that the area of the eardrum is only about $1\, cm^2$, so that only $10^{-16}\, W$ falls on it at the threshold of hearing! Air molecules in a sound wave of this intensity vibrate over a distance of less than one molecular diameter, and the gauge pressures involved are less than $10^{-9}\, atm$.

Table 5.2 – Sound Intensity Levels and Intensities

Sound intensity level β,dB	**Intensity** I,W/m^2	**Example/effect**
0	1×10^{-12}	Threshold of hearing
10	1×10^{-11}	Rustle of leaves
20	1×10^{-10}	Whisper at 1 m distance
30	1×10^{-9}	Quiet home
40	1×10^{-8}	Average home
50	1×10^{-7}	Average office, soft music
60	1×10^{-6}	Normal conversation
70	1×10^{-5}	Noisy office, busy traffic
80	1×10^{-4}	Loud radio, classroom lecture
90	1×10^{-3}	Inside a heavy truck
100	1×10^{-2}	Noisy factory
120	1	Threshold of pain
140	1×10^{2}	Jet airplane, severe pain
160	1×10^{4}	Bursting of eardrums

Sound intensity levels in decibels fit your experience better than intensities in watts per meter squared. The decibel scale is also easier to relate to because most people are more accustomed to dealing with numbers such as 0.53, or 120 than numbers such as 1.00×10^{-11}.

5.4. Hearing

The human ear has a tremendous range and sensitivity. It can give us a wealth of simple information – such as pitch, loudness, and direction.

Hearing is the perception of sound. (Perception is commonly defined to be awareness through the senses, a typically circular definition of higher level processes in living organisms.) Normal human hearing encompasses frequencies from 20 to $20,000$ Hz, an impressive range. Sounds below 20 Hz are called *infrasound*, whereas those above $20,000$ Hz are *ultrasound.* Neither is perceived by the ear, although infrasound can sometimes be felt as vibrations. When we do hear low-frequency vibrations, such as the sounds of a diving board, we hear the individual vibrations only because there are higher-frequency sounds in each. Other animals have hearing ranges different from that of humans. Dogs can hear sounds as high as $30,000$ Hz, whereas bats and dolphins can hear up to $100,000$ Hz sounds. You may have noticed that dogs respond to the sound of a dog whistle which produces sound out of the range of human hearing. Elephants are known to respond to frequencies below 20 Hz.

The perception of frequency is called *pitch.* Most of us have excellent relative pitch, which means that we can tell whether one sound has a different frequency from another. Typically, we can discriminate between two sounds if their frequencies differ by 0.3% or more. For example, 500.0 and 501.5 Hz are noticeably different. Pitch perception is directly related to frequency and is not greatly affected by other physical quantities such as intensity. Musical *notes* are particular sounds that can be produced by most instruments and in Western music have particular names. Combinations of notes constitute *music.* Some people can identify musical notes, such as A-sharp, C, or E-flat, just by listening to them. This uncommon ability is called perfect pitch.

The perception of intensity is called *loudness.* At a given frequency, it is possible to discern differences of about 1 dB, and a change of 3 dB is easily noticed. But loudness is not related to intensity alone. Frequency has a major effect on how loud a sound seems. The ear has its maximum sensitivity to frequencies in the range of 2000 to 5000 Hz, so that sounds in this range are perceived as being louder than, say, those at 500 or $10,000\ Hz$, even when they all have the same intensity. Sounds near the high- and low-frequency extremes of the hearing range seem even less loud, because the ear is even less sensitive at those frequencies.

Table 5.3 gives the dependence of certain human hearing perceptions on physical quantities.

Table 5.3 – Sound Perceptions

Perception	**Physical quantity**
Pitch	Frequency
Loudness	Intensity and Frequency
Timbre	Number and relative intensity of multiple frequencies. Subtle craftsmanship leads to non-linear effects and more details
Note	Basic unit of music with specific names, combined to generate tunes
Tone	Number and relative intensity of multiple frequencies

We call our perception of these combinations of frequencies and intensities *tone* quality, or more commonly the *timbre* of the sound. It is more difficult to correlate *timbre* perception to physical quantities than it is for loudness or pitch perception. Timbre is more subjective. Terms such as dull, brilliant, warm, cold, pure, and rich are employed to describe the timbre of a sound. So the consideration of timbre takes us into the realm of perceptual psychology, where higher-level processes in the brain are dominant.

This is true for other perceptions of sound, such as music and noise. We will concentrate on the question of loudness perception.

A unit called a *phon* is used to express loudness numerically. Phons differ from decibels because the phon is a unit of loudness perception, whereas the decibel is a unit of physical intensity. Figure 5.1 shows the relationship of loudness to intensity (or intensity level) and frequency for persons with normal hearing.

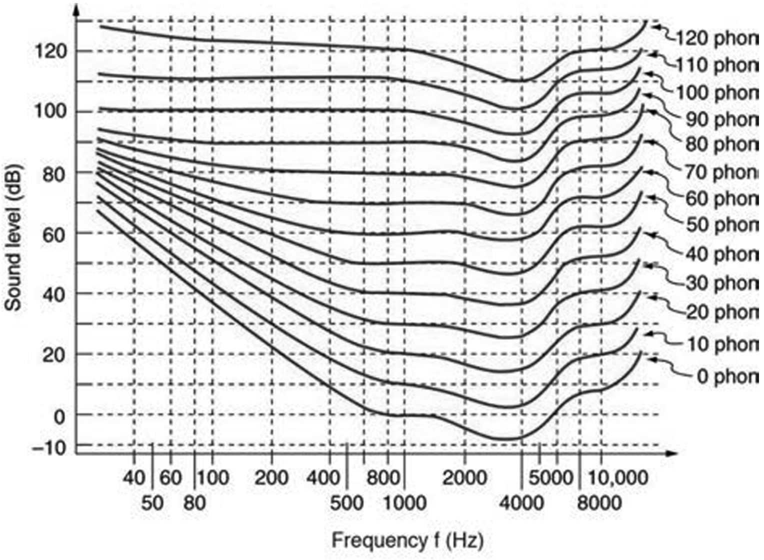

Figure 5.1 – The relationship of loudness in phons to intensity level (in decibels) and intensity (in watts per meter squared) for persons with normal hearing. The curved lines are equal-loudness curves – all sounds on a given curve are perceived as equally loud. Phons and decibels are defined to be the same at $1000 Hz$

The curved lines are equal-loudness curves. Each curve is labeled with its loudness in phons. Any sound along a given curve will be perceived as equally loud by the average person. The curves were determined by having large numbers of people compare the loudness of sounds at different frequencies and sound intensity levels. At a frequency of 1000 Hz, phons are taken to be numerically equal to decibels.

Further examination of the graph in Figure 5.1 reveals some interesting facts about human hearing. First, sounds below the 0-phon curve are not perceived by most people. So, for example, a 60 Hz sound at 40 dB is inaudible. The 0-phon curve represents the threshold of normal hearing. We can hear some sounds at intensity levels below 0 dB. For example, a 3 dB, 5000 Hz sound is audible, because it lies above the 0-phon curve. The loudness curves all have dips in them between about 2000 and 5000 Hz. These dips mean the ear is most sensitive to frequencies in that range.

The Hearing Mechanism

The hearing mechanism involves some interesting physics. The sound wave that impinges upon our ear is a pressure wave. The ear is a transducer that converts sound waves into electrical nerve impulses in a manner much more sophisticated than, but analogous to, a microphone. Figure 5.2 shows the gross anatomy of the ear with its division into three parts: the *outer ear* or ear canal; the *middle ear*, which runs from the *eardrum* to the *cochlea*; and the *inner ear*, which is the cochlea itself. The body part normally referred to as the ear is technically called the *pinna*.

The outer ear, or ear canal, carries sound to the recessed protected eardrum. The air column in the ear canal resonates and is partially responsible for the sensitivity of the ear to sounds in the 2000 to 5000 Hz range. The middle ear converts sound into

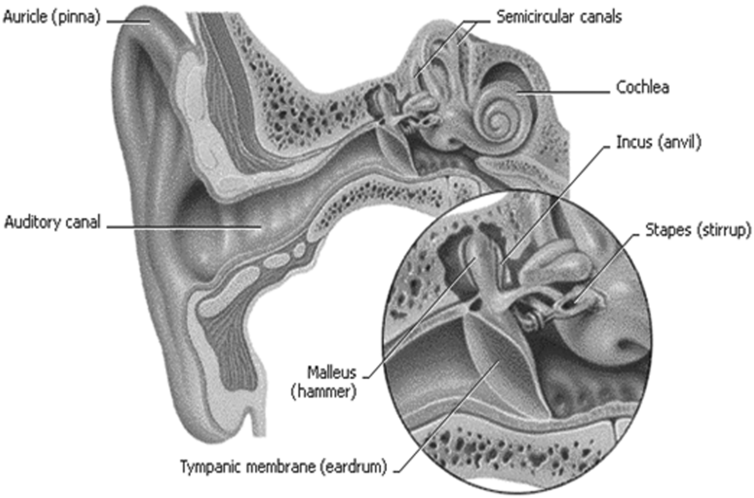

Figure 5.2 – The gross anatomy of the human ear

mechanical vibrations and applies these vibrations to the cochlea. The lever system of the middle ear takes the force exerted on the eardrum by sound pressure variations, amplifies it and transmits it to the inner ear via the oval window, creating pressure waves in the cochlea approximately 40 times greater than those impinging on the eardrum. (See Figure 5.3) Two muscles in the middle ear (not shown) protect the inner ear from very intense sounds. They react to intense sound in a few milliseconds and reduce the force transmitted to the cochlea. This protective reaction can also be triggered by your own voice, so that humming while shooting a gun, for example, can reduce noise damage.

Figure 5.4 shows the middle and inner ear in greater detail. Pressure waves moving through the cochlea cause the tectorial membrane to vibrate, rubbing cilia (called hair cells), which stimulate nerves that send electrical signals to the brain. The membrane resonates at different positions for different frequencies,

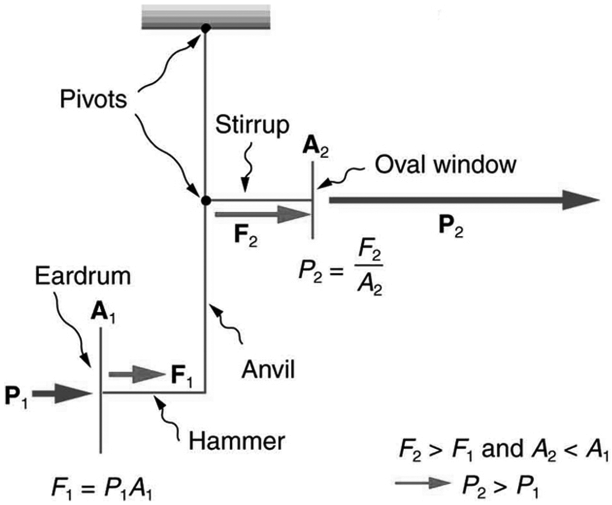

Figure 5.3 – This schematic shows the middle ear's system for converting sound pressure into force, increasing that force through a lever system, and applying the increased force to a small area of the cochlea

with high frequencies stimulating nerves at the near end and low frequencies at the far end. The complete operation of the cochlea is still not understood, but several mechanisms for sending information to the brain are known to be involved. For sounds below about $1000Hz$, the nerves send signals at the same frequency as the sound. For frequencies greater than about $1000\ Hz$, the nerves signal frequency by position. There is a structure to the cilia, and there are connections between nerve cells that perform signal processing before information is sent to the brain. Intensity information is partly indicated by the number of nerve signals and by volleys of signals. The brain processes the cochlear nerve signals to provide additional information such as source direction (based on time and intensity comparisons of sounds from both ears). Higher-level processing produces many nuances, such as music appreciation.

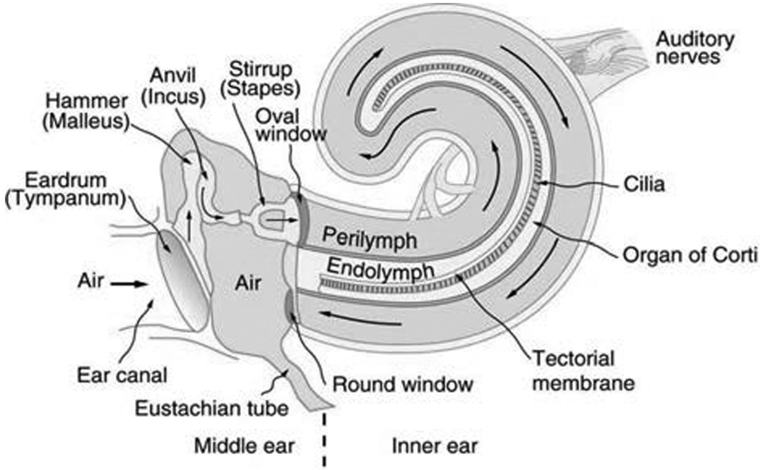

Figure 5.4 – The inner ear, or cochlea, is a coiled tube about 3 mm in diameter and 3 cm in length if uncoiled. When the oval window is forced inward, as shown, a pressure wave travels through the perilymph in the direction of the arrows, stimulating nerves at the base of cilia in the organ of Corti.

5.5. Doppler Effect

The*Doppler effect* is an alteration in the observed frequency of a sound due to motion of either the source or the observer. Although less familiar, this effect is easily noticed for a stationary source and moving observer. For example, if you ride a train past a stationary warning bell, you will hear the bell's frequency shift from high to low as you pass by. The actual change in frequency due to relative motion of source and observer is called a *Doppler shift*. The *Doppler effect* and *Doppler shift* are named for the Austrian physicist and mathematician Christian Johann Doppler (1803–1853), who did experiments with both moving sources and moving observers. The observer moving toward the source receives them at a higher frequency, and the person moving away from the source receives them at a lower frequency. A higher frequency is received by the observer moving toward the source, and a lower frequency is received by an observer moving away from the source. In general, then, relative motion of source and observer toward one another increases the received frequency. Relative motion apart decreases frequency. The greater the relative speed is, the greater the effect. For a stationary observer and a moving source, the frequency fobs received by the observer can be shown to be

$$f_{obs} = f_s \left(\frac{v}{v \pm v_s} \right), \tag{5.7}$$

where f_s is the frequency of the source, v_s is the speed of the source along a line joining the source and observer, and v is the speed of sound. The minus sign is used for motion toward the observer and the plus sign for motion away from the observer, producing the appropriate shifts up and down in frequency. Note that the greater the speed of the source, the greater the effect. Similarly, for a stationary source and moving observer, the frequency received by the observer fobs is given by

$$f_{obs} = f_s \left(\frac{v \pm v_{obs}}{v} \right), \tag{5.8}$$

where v_{obs} is the speed of the observer along a line joining the source and observer. Here the plus sign is for motion toward the source, and the minus is for motion away from the source.

The Doppler effect occurs not only for sound but for any wave when there is relative motion between the observer and the source.

There are Doppler shifts in the frequency of sound, light, and water waves, for example. Doppler shifts can be used to determine velocity, such as when ultrasound is reflected from blood in a medical diagnostic. The recession of galaxies is determined by the shift in the frequencies of light received from them and has implied much about the origins of the universe. Modern physics has been profoundly affected by observations of Doppler shifts.

5.6. Ultrasound

Any sound with a frequency above $20,000\ Hz$ (or $20\ kHz$) – that is, above the highest audible frequency – is defined to be *ultrasound*. In practice, it is possible to create ultrasound frequencies up to more than a gigahertz. (Higher frequencies are difficult to create; furthermore, they propagate poorly because they are very strongly absorbed.) Ultrasound has a tremendous number of applications, which range from burglar alarms to use in cleaning delicate objects to the guidance systems of bats. We begin our discussion of ultrasound with some of its applications in medicine, in which it is used extensively both for diagnosis and for therapy.

The characteristics of ultrasound, such as frequency and intensity, are wave properties common to all types of waves. Ultrasound, also, has a wavelength that limits the fineness of a detail it can detect. This characteristic is true of all waves. We can never observe details significantly smaller than the wavelength of our probe; for example, we will never see individual atoms with visible light, because the atoms are so small compared with the wavelength of light.

Ultrasound in Medical Therapy

Ultrasound, like any wave, carries energy that can be absorbed by the medium carrying it, producing effects that vary with intensity. When focused to intensities of 10^3 to 10^5 W/m^2, ultrasound can be used to shatter gallstones or pulverize cancerous tissue in surgical procedures. Great intensities can damage individual cells, variously causing their protoplasm to stream inside them, altering their permeability, or rupturing their walls through cavitation. Cavitation is the creation of vapor cavities in a fluid. The longitudinal vibrations in ultrasound alternatively compress and expand the medium, and at sufficient amplitudes the expansion separates molecules. Most cavitation damage is done when the cavities collapse, producing even greater shock pressures.

Most of the energy carried by high-intensity ultrasound in tissue is converted to thermal energy. In fact, intensities of 10^3 to 10^4 W/m^2 are commonly used for deep-heat treatments called *ultrasound diathermy*. Frequencies of 0.8 to 1 MHz are typical. In both athletics and physical therapy, ultrasound diathermy is most often applied to injured or overworked muscles to relieve pain and improve flexibility. Skill is needed by the therapist to avoid "bone burns" and other tissue damage caused by overheating and cavitation, sometimes made worse by reflection and focusing of the ultrasound by joint and bone tissue.

In some instances, you may encounter a different decibel scale, called the sound pressure level, when ultrasound travels in water or in human and other biological tissues. We shall not use the scale here, but it is notable that numbers for sound pressure levels range 60 to 70 dB higher than you would quote for β, the sound intensity level used in this text. Should you encounter a sound pressure level of 220 decibels, then, it is not an astronomically high intensity, but equivalent to about 155 dB – high enough to destroy tissue, but not as unreasonably high as it might seem at first.

Ultrasound in Medical Diagnostics

When used for imaging, ultrasonic waves are emitted from a transducer, a crystal exhibiting the piezoelectric effect (the expansion and contraction of a substance when a voltage is applied across it, causing a vibration of the crystal). These high-frequency vibrations are transmitted into any tissue in contact with the transducer. Similarly, if a pressure is applied to the crystal (in the form of a wave reflected off tissue layers), a voltage is produced which can be recorded. The crystal therefore acts as both a transmitter and a receiver of sound. Ultrasound is also partially absorbed by tissue on its path, both on its journey away from the transducer and on its return journey. From the time between when the original signal is sent and when the reflections from various boundaries between media are received, as well as a measure of the intensity loss of the signal, the nature and position of each boundary between tissues and organs may be deduced. Reflections at boundaries between two different media occur because of differences in a characteristic known as the *acoustic impedance* Z of each substance. Impedance is defined as

$$Z = \rho v, \tag{5.9}$$

where ρ is the density of the medium (in kg/m^3) and v is the speed of sound through the medium (in m/s). The units for Z are therefore $kg/(m^2 \cdot s)$.

Table 5.4 shows the density and speed of sound through various media (including various soft tissues) and the associated acoustic impedances. Note that the acoustic impedances for soft tissue do not vary much but there is a big difference between the acoustic impedance of soft tissue and air and also between soft tissue and bone.

At the boundary between media of different acoustic impedances, some of the wave energy is reflected and some is transmitted. The greater the difference in acoustic impedance between the two media, the greater the reflection and the smaller the transmission.

Table 5.4 – The Ultrasound Properties of Various Media, Including Soft Tissue Found in the Body

Medium	**Density** ρ, kg/m^3	**Speed of Ultrasound** v, m/s	**Acoustic Impedance** Z, $\times 10^6\ \frac{kg}{m^2 \cdot s}$
Air	1.3	330	429×10^{-6}
Water	1000	1500	1.5
Blood	1060	1570	1.66
Fat	925	1450	1.34
Muscle (average)	1075	1590	1.70
Bone (varies)	1400-1900	4080	5.7-7.8
Barium titanate (transducer material)	5600	5500	30.8

The intensity reflection coefficient α is defined as the ratio of the intensity of the reflected wave relative to the incident (transmitted) wave. This statement can be written mathematically as

$$\alpha = \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2}, \tag{5.10}$$

where Z_1 and Z_2 are the acoustic impedances of the two media making up the boundary. A reflection coefficient of zero (corresponding to total transmission and no reflection) occurs when the acoustic impedances of the two media are the same. An impedance "match" (no reflection) provides an efficient coupling of sound energy from one medium to another. The image formed in an ultrasound is made by tracking reflections and mapping the intensity of the reflected sound waves in a two-dimensional plane.

The applications of ultrasound in medical diagnostics have produced untold benefits with no known risks. Diagnostic intensities are too low (about 10^{-2} W/m^2) to cause thermal damage. More significantly, ultrasound has been in use for several decades and detailed follow-up studies do not show evidence of ill effects, quite unlike the case for X-rays.

The most common ultrasound applications produce an image like that shown in Figure 5.5. An ultrasound probe is held against the skin, with some coupling gel inbetween to help the ultrasound get into the body efficiently. A beam of ultrasound pulses, or μs tone bursts at MHz frequencies, are sent into the tissue. The wavelength of the sound in the tissue is typically 0.1-1 mm. The ultrasound waves propagate through the tissue and partially reflect, or scatter, from acoustic heterogeneities (differences in acoustic impedance), e.g., between different tissue types, fat, muscles, cysts, blood vessels, tumours, air pockets, etc. The reflected waves propagate back to the surface of the tissue. The reflected waves are detected at the surface, usually by the same transducer that sent them in, and recorded. The time of arrival of a pulse back at the detector indicates the depth from which the reflection came. Using this information, the recorded waves are processed to form an image. A computer constructs an image that reveals the shape and density of internal structures. At this stage, the image may also be processed to bring out certain features or to correct for the effect of acoustic absorption, for example. In clinical ultrasound imaging the images are captured in real-time and so single-frame snapshots give a poor impression of the amount of information in the images. The correct interpretation of US images depends on the skill and experience of the provider – who is, therefore, often a specially trained sonographer, radiographer or radiologist.

Figure 5.5 – Ultrasound image.

Doppler Ultrasound Instrumentation

Ultrasound is also in wide use to image the chambers of the heart and the flow of blood within the beating heart, using the Doppler effect (echocardiology). Moving targets (such as blood cells or a heart valve) with a velocity v_0 produce a Doppler shifted backscattered signal. The velocity is determined from the standard Doppler formula

$$v_0 = \frac{\Delta f}{2f\cos(\theta)} v, \tag{5.11}$$

where Δf is the frequency shift, f is the frequency of the transmitted wave ($f \gg \Delta f$), v is the sound velocity in the medium, and $\cos(\theta)$ is the angle between the wave axis and the velocity vector.

Lecture 6

Kinetic Theory and the Gas Laws

6.1. Temperature

The concept of temperature has evolved from the common concepts of hot and cold. Human perception of what feels hot or cold is a relative one. For example, if you place one hand in hot water and the other in cold water, and then place both hands in tepid water, the tepid water will feel cool to the hand that was in hot water, and warm to the one that was in cold water. The scientific definition of temperature is less ambiguous than your senses of hot and cold. *Temperature* is operationally defined to be what we measure with a thermometer. (Many physical quantities are defined solely in terms of how they are measured. We shall see later how temperature is related to the kinetic energies of atoms and molecules, a more physical explanation.)

The *Celsius* scale (which replaced the slightly different centigrade scale) has the freezing point of water at $0°C$ and the boiling point at $100°C$. Its unit is the degree Celsius ($°C$).

The *Kelvin* scale is the temperature scale that is commonly used in science. It is an absolute temperature scale defined to have 0 K at the lowest possible temperature, called *absolute zero*. The official temperature unit on this scale is the kelvin, which is abbreviated K, and is not accompanied by a degree sign. The freezing and boiling points of water are 273.15 K and 373.15 K, respectively. Thus, the magnitude of temperature differences is the same in units of kelvins and degrees Celsius. Unlike other temperature scales, the Kelvin scale is an absolute scale. It is used extensively in scientific work because a number of physical quantities, such as the volume of an ideal gas, are directly related to absolute temperature. The kelvin is the SI unit used in scientific work.

Temperatures on these scales can be converted using the equations in Table 6.1.

Table 6.1 – Temperature Conversions

To convert from	**Use this equation**
Celsius to Kelvin	$T(K) = T(°C) + 273.15$
Kelvin to Celsius	$T(°C) = T(K) - 273.15$

6.2. Thermal Equilibrium

Thermometers actually take their own temperature, not the temperature of the object they are measuring. This raises the question of how we can be certain that a thermometer measures the temperature of the object with which it is in contact. It is based on the fact that any two systems placed in thermal contact (meaning heat transfer can occur between them) will reach the same temperature. That is, heat will flow from the hotter object to the cooler one until they have exactly the same temperature. The objects are then in *thermal equilibrium*, and no further changes will occur.

Furthermore, experimentation has shown that if two systems, A and B, are in thermal equilibrium with each other, and B is in thermal equilibrium with a third system C, then A is also in thermal equilibrium with C. This conclusion may seem obvious, because all three have the same temperature, but it is basic to thermodynamics. It is called the *zeroth law of thermodynamics.*

6.3. Thermal Expansion of Solids and Liquids

An increase in temperature implies an increase in the kinetic energy of the individual atoms. In a solid, unlike in a gas, the atoms or molecules are closely packed together, but their kinetic energy (in the form of small, rapid vibrations) pushes neighboring atoms or molecules apart from each other. This neighbor-to-neighbor pushing results in a slightly greater distance, on average, between neighbors, and adds up to a larger size for the whole body. For most substances under ordinary conditions, there is no preferred direction, and an increase in temperature will increase the solid's size by a certain fraction in each dimension.

Linear Thermal Expansion

The change in length ΔL is proportional to length L. The dependence of thermal expansion on temperature, substance, and length is summarized in the equation

$$\Delta L = \alpha L \Delta T, \tag{6.1}$$

where ΔL is the change in length L, ΔT is the change in temperature, and α is the coefficient of linear expansion, which varies slightly with temperature.

Table 6.2 lists representative values of the coefficient of linear expansion, which may have units of $1/°C$ or $1/K$. Because the size of a kelvin and a degree Celsius are the same, both α and ΔTnn can be expressed in units of kelvins or degrees Celsius.

Thermal Expansion in Three Dimensions

The change in volume ΔV is very nearly $\Delta V = 3\alpha V \Delta T$. This equation is usually written as

$$\Delta V = \beta V \Delta T, \tag{6.2}$$

where β is the coefficient of volume expansion and $\beta \approx 3\alpha$.

Thermal Stress

Thermal stress is created by thermal expansion or contraction. Thermal stress can be destructive, such as when expanding gasoline ruptures a tank. It can also be useful, for example, when two parts are joined together by heating one in manufacturing, then slipping it over the other and allowing the combination to cool. Thermal stress can explain many phenomena, such as the weathering of rocks and pavement by the expansion of ice when it freezes.

Metal is regularly used in the human body for hip and knee implants. Most implants need to be replaced over time because, among other things, metal does not bond with bone. Researchers are trying to find better metal coatings that would allow metal-to-bone bonding. One challenge is to find a coating that has an expansion coefficient similar to that of metal. If the expansion coefficients are too different, the thermal stresses

Table 6.2 – Thermal Expansion Coefficients at $20 \circ C$

Material	**Coefficient of linear expansion** α, $(\times 10^{-6}\, 1/^{\circ}C)$	**Coefficient of volume expansion** β, $(\times 10^{-6}\, 1/^{\circ}C)$
Solids		
Aluminum	25	75
Brass	19	56
Copper	17	51
Gold	14	42
Iron	12	35
Invar	0.9	2.7
Lead	29	87
Silver	18	54
Glass	9	27
Quartz	0.4	1
Liquids		
Ether		1650
Ethyl alcohol		1100
Petrol		950
Glycerin		500
Mercury		180
Water		210
Gases		
Most other gases at atmospheric pressure		3400

during the manufacturing process lead to cracks at the coating-metal interface. Another example of thermal stress is found in the mouth. Dental fillings can expand differently from tooth enamel. It can give pain when eating ice cream or having a hot drink. Cracks might occur in the filling. Metal fillings (gold, silver, etc.)

are being replaced by composite fillings (porcelain), which have smaller coefficients of expansion, and are closer to those of teeth.

6.4. The Ideal Gas Law

In this section, we continue to explore the thermal behavior of gases. Because atoms and molecules in gases have large separations, forces between them can be ignored, except when they collide with each other during collisions. The motion of atoms and molecules (at temperatures well above the boiling temperature) is fast, such that the gas occupies all of the accessible volume and the expansion of gases is rapid. In contrast, in liquids and solids, atoms and molecules are closer together and are quite sensitive to the forces between them.

At room temperatures, collisions between atoms and molecules can be ignored. In this case, the gas is called an *ideal gas*, and the relationship between the pressure, volume, and temperature is given by the equation of state called the *ideal gas law.*

The *ideal gas* law states that

$$PV = NkT, \tag{6.3}$$

where P is the absolute pressure of a gas, V is the volume it occupies, N is the number of atoms and molecules in the gas, and T is its absolute temperature. The constant k is called the Boltzmann constant in honor of Austrian physicist Ludwig Boltzmann (1844-1906) and has the value

$$k = 1.38 \times 10^{-23}\ J/K.$$

Moles and Avogadro's Number

A *mole* (abbreviated mol) is defined to be the amount of a substance that contains as many atoms or molecules as there are atoms in exactly 12 *grams* (0.012 kg) of carbon-12. The actual number of atoms or molecules in one mole is called *Avogadro's number* (N_A), in recognition of Italian scientist Amedeo Avogadro (1776-1856). That is, the number is independent of the type of gas. This hypothesis has been confirmed, and the value of Avogadro's number is

$$N_A = 6.02 \times 10^{23}\, mol^{-1}.$$

A mole of any substance has a mass in grams equal to its *molecular mass*, which can be calculated from the atomic masses given in the periodic table of elements.

A very common expression of the ideal gas law uses the number of moles, n, rather than the number of atoms and molecules, N. Lets multiply and divide the equation (6.3) by Avogadro's number N_A. This gives

$$PV = \frac{N}{N_A} N_A kT.$$

Note that $n = N/N_A$ is the number of moles. We define the *universal gas constant* $R = N_A k$, and obtain the ideal gas law in terms of moles.

The ideal gas law (in terms of moles) is

$$PV = nRT. \tag{6.4}$$

The numerical value of R in SI units is $R = 8.31\, J/(mol \cdot K)$.

6.5. Kinetic Theory

We have developed macroscopic definitions of pressure and temperature. Pressure is the force divided by the area on which the force is exerted, and temperature is measured with a thermometer. We gain a better understanding of pressure and temperature from the kinetic theory of gases, which assumes that atoms and molecules are in continuous random motion.

Consider an elastic collision of a gas molecule with the wall of a container, so that it exerts a force on the wall (by Newton's third law). Because a huge number of molecules will collide with the wall in a short time, we observe an average force per unit area. These collisions are the source of pressure in a gas. As the number of molecules increases, the number of collisions and thus the pressure increase too. Similarly, the gas pressure is higher if the average velocity of molecules is higher. The following relationship is found:

$$PV = \frac{1}{3}Nm\overline{v^2}, \tag{6.5}$$

where P is the pressure (average force per unit area), V is the volume of gas in the container, N is the number of molecules in the container, m is the mass of a molecule, and $\overline{v^2}$ is the average molecular speed squared. Equating the right-hand side of equation (6.3) with the right-hand side of equation (6.3) gives

$$\frac{1}{3}Nm\overline{v^2} = NkT. \tag{6.6}$$

We can get the average kinetic energy of a molecule, $\frac{1}{2}m\overline{v^2}$, from the left-hand side of the equation (6.6) by canceling N and multiplying by 3/2. This calculation produces the result that the average kinetic energy of a molecule is directly related to absolute temperature.

$$\overline{E_K} = \frac{1}{2}m\overline{v^2} = \frac{3}{2}kT. \tag{6.7}$$

The average translational kinetic energy of a molecule, $\overline{E_K}$, is called *thermal energy.* The equation 6.7 is a molecular interpretation of temperature, and it has been found to be valid for gases and reasonably accurate in liquids and solids. It is another definition of temperature based on an expression of the molecular energy. It is sometimes useful to rearrange 6.7 and solve for the average speed of molecules in a gas in terms of temperature,

$$v_{rms} = \sqrt{\overline{v^2}} = \sqrt{\frac{3kT}{m}}, \tag{6.8}$$

where v_{rms} stands for root-mean-square (rms) speed.

Distribution of Molecular Speeds

The motion of molecules in a gas is random in magnitude and direction for individual molecules, but a gas of many molecules has a predictable distribution of molecular speeds. This distribution is called the *Maxwell-Boltzmann distribution*, after its originators, who calculated it based on kinetic theory, and has since been confirmed experimentally (see Figure 6.1). The distribution has a long tail, because a few molecules may go several times the rms speed. The most probable speed v_p is less than the rms speed v_{rms}. The distribution of thermal speeds depends strongly on temperature. As temperature increases, the speeds are shifted to higher values and the distribution is broadened.

Phase Changes

Up to now, we have considered the behavior of ideal gases. Real gases are like ideal gases at high temperatures. At lower temperatures, however, the interactions between the molecules and their volumes cannot be ignored. The molecules are very close (condensation occurs), and there is a dramatic decrease in volume. The substance changes from a gas to a liquid. When a liquid is cooled to even lower temperatures, it becomes a solid. The volume never reaches zero because of the finite volume of the molecules.

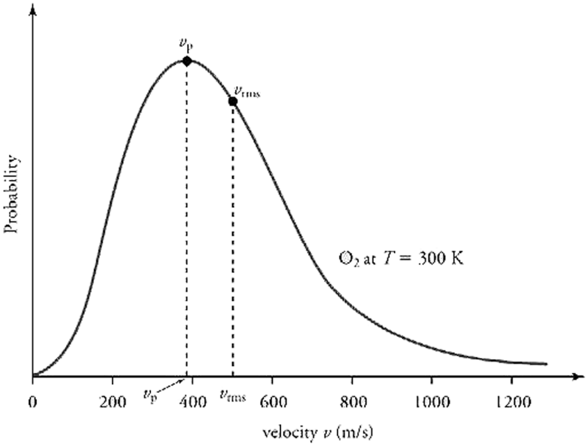

Figure 6.1 – The Maxwell-Boltzmann distribution of molecular speeds in an ideal gas

6.6. Heat Transfer

Energy can exist in many forms and heat is one of the most intriguing. Heat is often hidden, as it only exists when in transit, and is transferred by a number of distinctly different methods. Heat transfer touches every aspect of our lives and helps us understand how the universe functions. It explains the chill we feel on a clear breezy night, or why Earth's core has yet to cool.

Heat

Work is defined as force times distance and learned that work done on an object changes its kinetic energy. We also saw that temperature is proportional to the (average) kinetic energy of atoms and molecules. We say that a thermal system has a certain internal energy: its internal energy is higher if the temperature is higher. If two objects at different temperatures are brought in contact with each other, energy is transferred from the hotter to the colder object until equilibrium is reached and the bodies reach thermal equilibrium (i.e., they are at the same temperature). No work is done by either object, because no force acts through a distance. The transfer of energy is caused by the temperature difference and ceases once the temperatures are equal. These observations lead to the following definition of *heat*: *Heat is the spontaneous transfer of energy due to a temperature difference.*

Heat is often confused with temperature. For example, we may say the heat was unbearable, when we actually mean that the temperature was high. Heat is a form of energy, whereas temperature is not. The misconception arises because we are sensitive to the flow of heat, rather than the temperature.

Owing to the fact that heat is a form of energy, it has the SI unit of joule (J). The calorie (cal) is a common unit of energy, defined as the energy needed to change the temperature of 1.00 g of water by $1.00°C$ – specifically, between $14.5°C$ and $15.5°C$, since there is a slight temperature dependence. Perhaps the most common unit of heat is the kilocalorie ($kcal$), which is the energy needed to change the temperature of 1.00 kg of water by $1.00°C$. Since mass is most often specified in kilograms, kilocalorie is commonly used. Food calories (given the notation Cal, and sometimes called "big calorie") are actually kilocalories ($1\, kilocalorie = 1000\, calories$), a fact not easily determined from package labeling.

Mechanical Equivalent of Heat

It is also possible to change the temperature of a substance by doing work. Work can transfer energy into or out of a system. This realization helped establish the fact that heat is a form of energy. James Prescott Joule (1818-1889) performed many experiments to establish the mechanical equivalent of heat – *the work needed to* *produce the same effects as heat transfer.* In terms of the units used for these two terms, the best modern value for this equivalence is

$$1.000\,kcal = 4186\,J.$$

Temperature Change and Heat Capacity

One of the major effects of heat transfer is temperature change: heating increases the temperature while cooling decreases it. We assume that there is no phase change and that no work is done on or by the system. Experiments show that the transferred heat depends on three factors: the change in temperature, the mass of the system, and the substance, and phase of the substance.

The quantitative relationship between heat transfer and temperature change contains all three factors:

$$Q = mc\Delta T, \tag{6.9}$$

where Q is the symbol for heat transfer, m is the mass of the substance, and ΔT is the change in temperature. The symbol c stands for *specific heat* and depends on the material and phase. The *specific heat* is the amount of heat necessary to change the temperature of $1.00\,kg$ of mass by $1.00°C$. The specific heat c is a property of the substance; its SI unit is $J/(kg \cdot K)$ or $J/(kg \cdot^\circ C)$. Recall that the temperature change (ΔT) is the same in units of kelvin and degrees Celsius. If heat transfer is measured in kilocalories, then the unit of specific heat is $kcal/(kg \cdot^\circ C)$.

Values of specific heat must generally be looked up in tables, because there is no simple way to calculate them. In general, the specific heat also depends on the temperature. Table 6.3 lists representative values of specific heat for various substances. Except for gases, the temperature and volume dependence of the specific heat of most substances is weak.

Table 6.3 – Specific Heats of Various Substances

Substances	**Specific heat, c**	
	$J/(kg \cdot^\circ C)$	$kcal/(kg \cdot^\circ C)$
Solids		
Aluminum	900	0.215
Copper	387	0.0924
Iron	452	0.108
Silver	235	0.0562
Glass	840	0.20
Human body (average)	3500	0.83
Ice (average)	2090	0.50
Wood	1700	0.4
Liquids		
Ethanol	2450	0.586
Glycerin	2410	0.576
Mercury	139	0.0333
Water (15.0°C)	4186	1.000
Gases[1]		
Air (dry)	721 (1015)	0.172 (0.242)
Ammonia	1670 (2190)	0.399 (0.523)
Carbon dioxide	638 (833)	0.152 (0.199)
Nitrogen	739 (1040)	0.177 (0.248)
Oxygen	651 (913)	0.156 (0.218)
Steam (100°C)	1520 (2020)	0.363 (0.482)

---

[1] c_v at constant volume and at 20.0°C, except as noted, and at 1.00 atm average pressure. Values in parentheses are c_p at a constant pressure of 1.00 atm.

Phase Change and Latent Heat

The energy involved in a phase change depends on two major factors: the number and strength of bonds or force pairs. The number of bonds is proportional to the number of molecules and thus to the mass of the sample. The strength of forces depends on the type of molecules. The heat Q required to change the phase of a sample of mass m is given by

$$Q = mL_f \qquad \text{(melting/freezing)}, \tag{6.10}$$

$$Q = mL_v \qquad \text{(vaporization/condensation)}, \tag{6.11}$$

where the *latent heat of fusion*, L_f, and *latent heat of vaporization*, L_v, are material constants that are determined experimentally. See (Table 6.4).

Latent heat is measured in units of J/kg. Both L_f and L_v depend on the substance, particularly on the strength of its molecular forces as noted earlier. L_f and L_v are collectively called *latent heat coefficients*. They are *latent*, or hidden, because in phase changes, energy enters or leaves a system without causing a temperature change in the system; so, in effect, the energy is hidden. Table 6.4 lists representative values of L_f and L_v, together with melting and boiling points.

Table 6.4 – Heats of Fusion and Vaporization

Substances	**Melting point** (°C)	L_f		**Boiling point** (°C)	L_v	
		kJ/kg	$kcal/kg$		kJ/kg	$kcal/kg$
Helium	−269.7	5.23	1.25	−268.9	20.9	4.99
Hydrogen	−259.3	58.6	14.0	−252.9	452	108
Nitrogen	−210.0	25.5	6.09	−195.8	201	48.0
Oxygen	−218.8	13.8	3.30	−183.0	213	50.9
Ethanol	−114	104	24.9	78.3	854	204
Ammonia	−75	452.5	108	−33.4	1370	327
Mercury	−38.9	11.8	2.82	357	272	65.0
Water	0.00	334	79.8	100.0	2256	539
Sulfur	119	38.1	9.10	444.6	326	77.9
Lead	327	24.5	5.85	1750	871	208
Antimony	631	165	39.4	1440	561	134
Aluminum	660	380	90	2450	11400	2720
Silver	961	88.3	21.1	2193	2336	558
Gold	1063	64.5	15.4	2660	1578	377
Copper	1083	134	32.0	2595	5069	1211
Uranium	1133	84	207	3900	1900	454
Tungsten	3410	184	44	5900	4810	1150

We examine the effects of phase change more precisely by considering adding heat into a sample of ice at $-20°C$ (Figure 6.2).

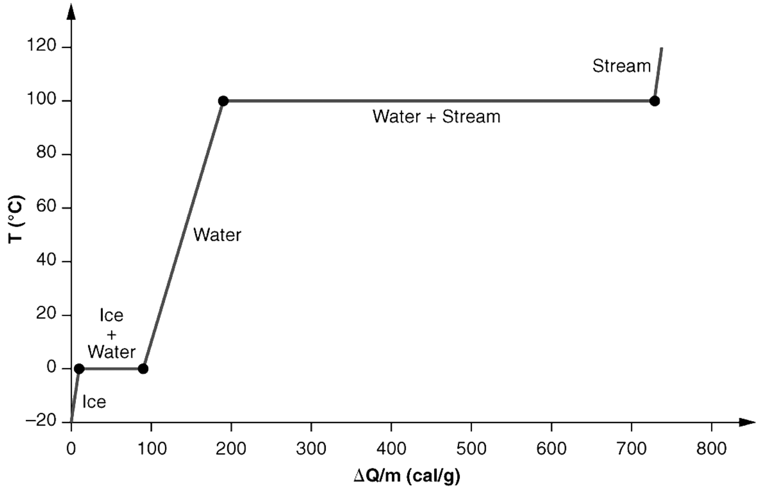

Figure 6.2 – A graph of temperature versus energy added. The system is constructed so that no vapor evaporates while ice warms to become liquid water, and so that, when vaporization occurs, the vapor remains in the system. The long stretches of constant temperature values at $0°C$ and $100°C$ reflect the large latent heat of melting and vaporization, respectively

The temperature of the ice rises linearly, absorbing heat at a constant rate of $0.50\, cal/g\cdot°C$ until it reaches $0°C$. Once at this temperature the ice begins to melt until all the ice has melted, absorbing $79.8\, cal/g$ of heat, the temperature remains constant at $0°C$ during this phase change. Once all the ice has melted, the temperature of the liquid water rises, absorbing heat at a new constant rate of $1.00 cal/g \cdot^{\circ} C$. At $100°C$, the water begins to boil and the temperature again remains constant while the water absorbs 539 cal/g of heat during this phase change. When all the liquid has become steam vapor, the temperature rises again, absorbing heat at a rate of 0.482 $cal/g \cdot^{\circ} C$.

Water can evaporate at temperatures below the boiling point. More energy is required than at the boiling point, because the kinetic energy of water molecules at temperatures below $100°C$ is less than that at $100°C$, hence less energy is available from random thermal motions. Take, for example, the fact that, at body temperature, perspiration from the skin requires a heat input of 2428 kJ/kg, which is about 10 percent higher than the latent heat of vaporization at $100°C$. This heat comes from the skin and thus provides an effective cooling mechanism in hot weather. High humidity inhibits evaporation so that body temperature might rise, leaving unevaporated sweat on your brow.

6.7. Heat Transfer Methods

Equally as interesting as the effects of heat transfer on a system are the methods by which this occurs. Whenever there is a temperature difference, heat transfer occurs. Heat transfer may occur rapidly, such as through a cooking pan, or slowly, such as through the walls of a picnic ice chest. We can control rates of heat transfer by choosing materials (such as thick wool clothing for the winter), controlling air movement (such as the use of weather stripping around doors), or by choice of color (such as a white roof to reflect summer sunlight). So many processes involve heat transfer, so that it is hard to imagine a situation where no heat transfer occurs. Yet every process involving heat transfer takes place by only three methods:

1. Conduction is heat transfer through stationary matter by physical contact. (The matter is stationary on a macroscopic

scale – we know there is thermal motion of the atoms and molecules at any temperature above absolute zero.) Heat transferred between the electric burner of a stove and the bottom of a pan is transferred by conduction.

2. Convection is the heat transfer by the macroscopic movement of a fluid. This type of transfer takes place in a forced-air furnace and in weather systems, for example.
3. Heat transfer by radiation occurs when microwaves, infrared radiation, visible light, or another form of electromagnetic radiation is emitted or absorbed. An obvious example is the warming of the Earth by the Sun. A less obvious example is thermal radiation from the human body.

Conduction

Lastly, the heat transfer rate depends on the material properties described by the coefficient of thermal conductivity. The *rate of conductive heat transfer* through a slab of material is given by

$$\frac{Q}{t} = \frac{kA(T_2 - T_1)}{d}, \tag{6.12}$$

where Q/t is the rate of heat transfer in watts or kilocalories per second, k is the thermal conductivity of the material, A and d are its surface area and thickness, and $(T_2 - T_1)$ is the temperature difference across the slab. Table 6.5 gives representative values of thermal conductivity.

A combination of material and thickness is often manipulated to develop good insulators – the smaller the conductivity k and the larger the thickness d, the better. The ratio of d/k will thus be large for a good insulator. The ratio d/k is called the R factor. The rate of conductive heat transfer is inversely proportional to R. The larger the value of R, the better the insulation.

Table 6.5 – Thermal Conductivities of Common Substances

Substance	**Thermal conductivity** k, $J/(s \cdot m \cdot^{\circ} C)$
Silver	420
Copper	390
Gold	318
Aluminum	220
Steel	80
Ice	2.2
Glass (average)	0.84
Water	0.6
Fatty tissue (without blood)	0.2
Asbestos	0.16
Plasterboard	0.16
Wood	0.08-0.16
Snow (dry)	0.10
Cork	0.042
Glass wool	0.042
Wool	0.04
Down feathers	0.025
Air	0.023
Styrofoam	0.010

Convection

Convection is driven by large-scale flow of matter. In the case of Earth, the atmospheric circulation is caused by the flow of hot air from the tropics to the poles, and the flow of cold air from the poles toward the tropics. The circulatory system is used in the body: when the body overheats, the blood vessels in the skin expand (dilate), which increases the blood flow to the skin where it can be cooled by sweating. These vessels become smaller when it is cold outside and larger when it is hot (so more fluid flows, more energy is transferred).

The body also loses a significant fraction of its heat through the breathing process.

While convection is usually more complicated than conduction, we can describe convection and do some straightforward, realistic calculations of its effects. Natural convection is driven by buoyant forces: hot air rises because density decreases as temperature increases.

Some interesting phenomena happen when convection is accompanied by a phase change. It allows us to cool off by sweating, even if the temperature of the surrounding air exceeds body temperature. Heat from the skin is required for sweat to evaporate from the skin, but without air flow, the air becomes saturated and evaporation stops. Air flow caused by convection replaces the saturated air by dry air and evaporation continues.

Another important example of the combination of phase change and convection occurs when water evaporates from the oceans. Heat is removed from the ocean when water evaporates. If the water vapor condenses in liquid droplets as clouds form, heat is released in the atmosphere. Thus, there is an overall transfer of heat from the ocean to the atmosphere.

Radiation

Heat is transferred by radiation. That is, the hot body emits electromagnetic waves that are absorbed by our skin: no medium is required for electromagnetic waves to propagate. Different names are used for electromagnetic waves of different wavelengths: radio waves, microwaves, infrared radiation, visible light, ultraviolet radiation, X-rays, and gamma rays.

The energy of electromagnetic radiation depends on the wavelength (color) and varies over a wide range: a smaller wavelength (or higher frequency) corresponds to a higher energy. Because more heat is radiated at higher temperatures, a temperature change is accompanied by a color change. The radiated energy depends on its intensity, which is represented in the Figure 6.3 by the height of the distribution.

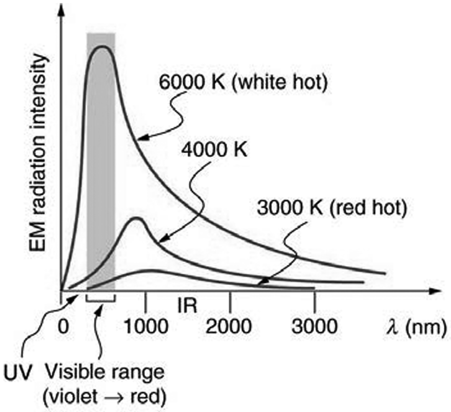

Figure 6.3 – A graph of the spectra of electromagnetic waves emitted from an ideal radiator at three different temperatures

All objects absorb and emit electromagnetic radiation. The rate of heat transfer by radiation is largely determined by the color of the object. Black is the most effective, and white is the least effective. An *ideal radiator* is the same color as an *ideal absorber* and captures all the radiation that falls on it. In contrast, white is a poor absorber and is also a poor radiator. A white object reflects all radiation, like a mirror.

The rate of heat transfer by emitted radiation is determined by the Stefan-Boltzmann law of radiation:

$$\frac{Q}{t} = \sigma\alpha A T^4, \tag{6.13}$$

where $\sigma = 5.67 \times 10^{-8}\ J/(s \cdot m^2 \cdot K^4)$ is the Stefan-Boltzmann constant, A is the surface area of the object, and T is its absolute temperature in kelvin. The symbol α stands for the emissivity of the object, which is a measure of how well it radiates. An ideal jet-black (or *black body*) radiator has $\alpha = 1$, whereas a perfect reflector has $\alpha = 0$. Real objects fall between these two values. Take, for example, tungsten light bulb filaments which have an e of about 0.5, and carbon black (a material used in printer toner), which has the (greatest known) emissivity of about 0.99.

Skin is a remarkably good absorber and emitter of infrared radiation, having an emissivity of 0.97 in the infrared spectrum. Thus, we are all nearly (jet) black in the infrared, in spite of the obvious variations in skin color. This high infrared emissivity is why we can so easily feel radiation on our skin. It is also the basis for the use of night scopes used by law enforcement and the military to detect human beings. Even small temperature variations can be detected because of the T^4 dependence. Images, called thermographs, can be used medically to detect regions of abnormally high temperature in the body, perhaps indicative of disease.

All objects emit and absorb radiation. The net rate of heat transfer by radiation (absorption minus emission) is related to both the temperature of the object and the temperature of its surroundings. Assuming that an object with a temperature T_1 is surrounded by an environment with uniform temperature T_2 , the net rate of heat transfer by radiation is

$$\frac{Q_{net}}{t} = \sigma \alpha A \left(T_2^4 - T_1^4 \right), \tag{6.14}$$

where α is the emissivity of the object alone. In other words, it does not matter whether the surroundings are white, gray, or black; the balance of radiation into and out of the object depends on how well it emits and absorbs radiation. When $T_2 > T_1$, the quantity Q_{net}/t is positive – that is, the net heat transfer is from hot to cold.

Problem-Solving Strategies for the Methods of Heat Transfer

1. Examine the situation to determine what type of heat transfer is involved.
2. Identify the type(s) of heat transfer – conduction, convection, or radiation.
3. Identify exactly what needs to be determined in the problem (identify the unknowns). A written list is very useful.
4. Make a list of what is given or can be inferred from the problem as stated (identify the knowns).
5. Solve the appropriate equation for the quantity to be determined (the unknown).
6. For conduction, equation $\frac{Q}{t} = \frac{kA(T_2 - T_1)}{d}$ is appropriate. Table 6.5 lists thermal conductivities. For convection, determine the amount of matter moved and use equation $Q = mc\Delta T$ to calculate the heat transfer involved in the temperature change of the fluid. If a phase change accompanies convection, equation $Q = mL_f$ or $Q = mL_v$ is appropriate to find the heat transfer involved in the phase

change. Table 6.4 lists information relevant to the phase change. For radiation, equation $\frac{Q_{net}}{t} = \sigma\alpha A\left(T_2^4 - T_1^4\right)$ gives the net heat transfer rate.

7. Insert the knowns along with their units into the appropriate equation and obtain numerical solutions complete with units.
8. Check the answer to see if it is reasonable. Does it make sense?

Lecture 7

Thermodynamics

7.1. The First Law of Thermodynamics

If we are interested in how heat transfer is converted into doing work, then the conservation of energy principle is important. The first law of thermodynamics applies the conservation of energy principle to systems where heat transfer and doing work are the methods of transferring energy into and out of the system. The *first law of thermodynamics* states that the change in internal energy of a system equals the net heat transfer into the system minus the net work done by the system. In equation form, the first law of thermodynamics is

$$\Delta U = Q - W. \tag{7.1}$$

Here ΔU is the change in internal energy U of the system. Q is the net heat transferred into the system – that is, Q is the sum of all heat transfer into and out of the system. W is the net work done by the system – that is, W is the sum of all work done on or by the system. We use the following sign conventions: if Q is positive, then there is a net heat transfer into the system; if W is positive, then there is net work done by the system. So positive Q

adds energy to the system and positive W takes energy from the system. Thus $\Delta U = Q - W$. Note, also, that if more heat transfer into the system occurs than work done, the difference is stored as internal energy. Heat engines are a good example of this – heat transfer into them takes place so that they can do work.

Internal Energy U

We can think about the internal energy of a system in two different but consistent ways. The first is the atomic and molecular view, which examines the system on the atomic and molecular scale. The *internal energy U* of a system is the sum of the kinetic and potential energies of its atoms and molecules. Recall that kinetic plus potential energy is called mechanical energy. Thus internal energy is the sum of atomic and molecular mechanical energy. Because it is impossible to keep track of all individual atoms and molecules, we must deal with averages and distributions. A second way to view the internal energy of a system is in terms of its macroscopic characteristics, which are very similar to atomic and molecular average values. Macroscopically, we define the change in internal energy ΔU to be that given by the first law of thermodynamics (7.1). Many detailed experiments have verified that ΔU is the change in total kinetic and potential energy of all atoms and molecules in a system. It has also been determined experimentally that the internal energy U of a system depends only on the state of the system and not how it reached that state. More specifically, U is found to be a function of a few macroscopic quantities (pressure, volume, and temperature, for example), independent of past history such as whether, there has been heat transfer or work done. This independence means that if we know the state of a system, we can calculate changes in its internal energy U from a few macroscopic variables.

To get a better idea of how to think about the internal energy of a system, let us examine a system going from state 1 to state 2. The system has internal energy U_1 in state 1, and it has internal energy U_2 in state 2, no matter how it got to either state. So the change in internal energy $\Delta U = U_2 - U_1$ is independent of what caused the change. In other words, ΔU is *independent of path.* By path, we mean the method of getting from the starting point to the ending point.

For the special case of a *monatomic ideal gas*, we note that the average kinetic energy of an atom in such a gas is given by

$$\frac{1}{2}m\overline{v^2} = \frac{3}{2}kT.$$

The kinetic energy of the atoms in a monatomic ideal gas is its only form of internal energy, and so its total internal energy U is

$$U = N\frac{1}{2}m\overline{v^2} = \frac{3}{2}NkT = \frac{3}{2}nRT, \qquad \text{(monatomic ideal gas)}, \tag{7.2}$$

where N is the number of atoms in the gas, n is the number of moles.

This relationship means that the internal energy of an ideal monatomic gas is constant during an *isothermal process* – that is, $\Delta U = 0$. If the internal energy does not change, then the net heat transfer into the gas must equal the net work done by the gas. That is, because $\Delta U = Q - W = 0$ here, $Q = W$.

An *adiabatic process* is defined to be one in which there is no heat transfer – that is, $Q = 0$. Processes that are nearly adiabatic can be achieved either by using very effective insulation or by performing the process so fast that there is little time for heat transfer. In fact, because $Q = 0$, $\Delta U = -W$ for an adiabatic process.

Heat Q and Work W

Heat transfer Q and doing work W are the two everyday means of bringing energy into or taking energy out of a system. The processes are quite different. Heat transfer, a less organized process, is driven by temperature differences. Work, a quite organized process, involves a macroscopic force exerted through a distance. Nevertheless, heat and work can produce identical results. For example, both can cause a temperature increase. Heat transfer into a system, such as when the Sun warms the air in a bicycle tire, can increase its temperature, and so can work done on the system, as when the bicyclist pumps air into the tire. Once the temperature increase has occurred, it is impossible to tell whether it was caused by heat transfer or by doing work. This uncertainty is an important point. Heat transfer and work are both energy in transit – either is stored as such in a system. However, both can change the internal energy U of a system. Internal energy is a form of energy completely different from either heat or work.

Work Done on or by a Gas

A process by which a gas does work on a piston at constant pressure is called an *isobaric process* (see Figure 7.1). Since the pressure is constant, the force exerted is constant and the work done is given as

$$W = Fd,$$

where F is force acting on the piston, d is its displacement.

From definition of pressure follows that $F = PA$, and so

$$W = PAd.$$

Because the volume of a cylinder is its cross-sectional area A times its length, we see that $Ad = \Delta V$, the change in volume; thus,

$$W = P\Delta V \qquad \text{(isobaric process)}. \tag{7.3}$$

Note that if ΔV is positive, then W is positive, meaning that work is done by the gas on the outside world.

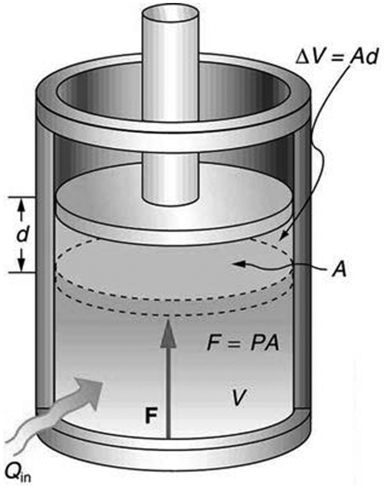

Figure 7.1 – An isobaric expansion of a gas requires heat transfer to keep the pressure constant

Processes when volume is constant are called *isochoric processes.* Since volume is constant, $\Delta V = 0$, and no work is done in an isochoric process.

In general case, work done on or by gas is determined by formula:

$$W = \int_{V_1}^{V_2} P dV. \tag{7.4}$$

Figure 7.2 shows a graph of pressure versus volume – that is, a PV diagram for an isobaric process. You can see in the figure that the work done is the area under the graph. This property of

PV diagrams is very useful and broadly applicable: *the work done on or by a system in going from one state to another equals the area under the curve on a PV diagram.*

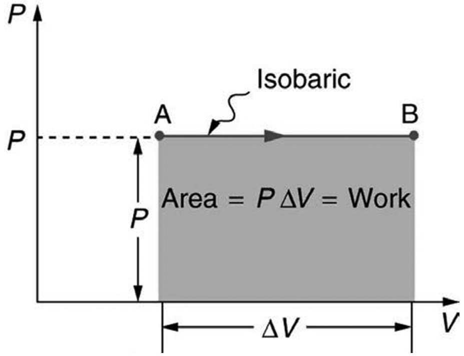

Figure 7.2 – A graph of pressure versus volume for a constant-pressure, or isobaric, process. The area under the curve equals the work done by the gas, since $W = P\Delta V$.

Thus, the total work done is the total area under the curve. If the path is reversed, then work is done on the system. The area under the curve in that case is negative, because ΔV is negative. *PV* diagrams clearly illustrate that *the work done depends on the path taken and not just the endpoints.* If the system follows the cyclical path, as in Figure 7.3, then the total work done is the area inside the loop.

Reversible Processes

A *reversible process* is one in which both the system and its environment can return to exactly the states they were in by following the reverse path. Real macroscopic processes are never exactly reversible. Reversibility requires the direction of

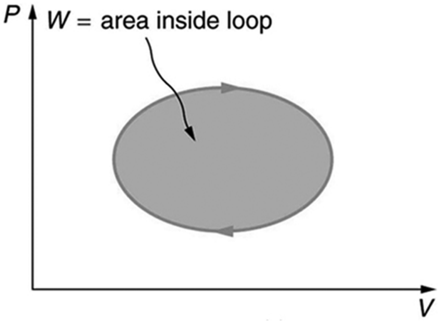

Figure 7.3 – The area inside any closed loop is the work done in the cyclical process. If the loop is traversed in a clockwise direction, W is positive – it is work done on the outside environment. If the loop is traveled in a counter-clockwise direction, W is negative – it is work that is done to the system

heat transfer to reverse for the reverse path. Since dissipative mechanisms cannot be completely eliminated, real processes cannot be reversible.

Human Metabolism and the First Law of Thermodynamics

Human metabolism is the conversion of food into heat transfer, work, and stored fat. Metabolism is an interesting example of the first law of thermodynamics in action. We now take another look at these topics via the first law of thermodynamics. Considering the body as the system of interest, we can use the first law to examine heat transfer, doing work, and internal energy in activities ranging from sleep to heavy exercise. What are some of the major characteristics of heat transfer, doing work, and energy in the body? First, body temperature is normally kept constant by heat transfer to the surroundings. This means that Q is negative. Another fact is that the body, usually, does work on the outside world. This means that W is positive. In such situations, then, the body loses internal energy, since $\Delta U = Q - W < 0$ or is negative.

Now consider the effects of eating. Eating increases the internal energy of the body by adding chemical potential energy (this is an unromantic view of a good steak). The body metabolizes all the food we consume. Basically, metabolism is an oxidation process in which the chemical potential energy of food is released. This implies that food input is in the form of work. Food energy is reported in a special unit, known as the *Calorie.* This energy is measured by burning food in a calorimeter, which is how the units are determined.

In chemistry and biochemistry, one *calorie* (spelled with a lowercase *c*) is defined as the energy (or heat transfer) required to raise the temperature of one gram of pure water by one degree Celsius. Nutritionists and weight-watchers tend to use the dietary calorie, which is frequently called a *Calorie* (spelled with a capital *C*). One food *Calorie* is the energy needed to raise the temperature of one kilogram of water by one degree Celsius. This means that one dietary *Calorie* is equal to one kilocalorie for the chemist, and one must be careful to avoid confusion between the two.

Again, consider the internal energy the body has lost. There are three places this internal energy can go – to heat transfer, to doing work, and to stored fat (a tiny fraction also goes to cell repair and growth). Heat transfer and doing work take internal energy out of the body, and food puts it back. If you eat just the right amount of food, then your average internal energy remains constant. Whatever you lose to heat transfer and doing work is replaced by food, so that, in the long run, $\Delta U = 0$. If you overeat repeatedly, then ΔU is always positive, and your body stores this extra internal energy as fat. The reverse is true if you eat too little. If ΔU is negative for a few days, then the body metabolizes its own fat to maintain body temperature and do work that takes energy from the body. This process is how dieting produces weight loss.

The body provides us with an excellent indication that many thermodynamic processes are irreversible. An *irreversible process* can go in one direction but not the reverse, under a given set of conditions. For example, although body fat can be converted to do work and produce heat transfer, work done on the body and heat transfer into it cannot be converted to body fat.

7.2. The Second Law of Thermodynamics

The fact that certain processes never occur suggests that there is a law forbidding them to occur. The first law of thermodynamics would allow them to occur – none of those processes violate conservation of energy. The law that forbids these processes is called the second law of thermodynamics. We shall see that the second law can be stated in many ways that may seem different, but which in fact are equivalent. Like all natural laws, the second law of thermodynamics gives insights into nature, and its several statements imply that it is broadly applicable, fundamentally affecting many apparently disparate processes.

The Second Law of Thermodynamics (first expression) *Heat transfer occurs spontaneously from higher- to lower-temperature bodies but never spontaneously in the reverse direction.* Another way of stating this: It is impossible for any process to have as its sole result heat transfer from a cooler to a hotter object.

Heat Engine

One of the most important things we can do with heat transfer is to use it to do work for us. Such a device is called a *heat engine*, and one is shown schematically in Figure 7.4. Gasoline and diesel engines, jet engines, and steam turbines are all heat engines that do work by using part of the heat transfer from some source. Heat transfer from the hot object (or hot reservoir) is denoted as Q_h , while heat transfer into the cold object (or cold reservoir) is Q_c, and the work done by the engine is W. The temperatures of the hot and cold reservoirs are T_h and T_c, respectively.

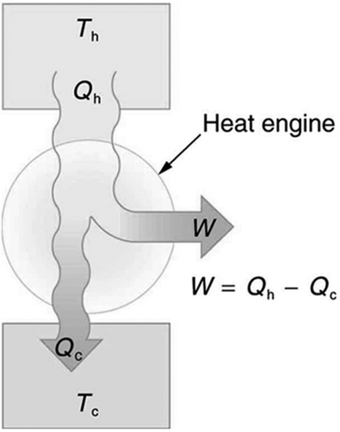

Figure 7.4 – A heat engine, represented here by a circle, uses part of the heat transfer to do work

Because the hot reservoir is heated externally, which is energy intensive, it is important that the work is done as efficiently as possible. In fact, we would like W to equal Q_h, and for there to be no heat transfer to the environment ($Q_c = 0$). Unfortunately, this is impossible.

The Second Law of Thermodynamics (second expression) *It is impossible in any system for heat transfer from a reservoir to completely convert to work in a cyclical process in which the system returns to its initial state.*

In the conversion of energy to work, we are always faced with the problem of getting less out than we put in. We define conversion efficiency E_{ff} to be the ratio of useful work output to the energy input (or, in other words, the ratio of what we get to what we spend). In that spirit, we define the efficiency of a heat engine to be its net work output W divided by heat transfer to the engine Q_h – that is,

$$E_{ff} = \frac{W}{Q_h}.$$

Since $W = Q_h - Q_c$ in a cyclical process, we can also express this as

$$E_{ff} = \frac{Q_h - Q_c}{Q_h} = 1 - \frac{Q_c}{Q_h}. \qquad \text{(cyclical process)} \tag{7.5}$$

Carnot's Perfect Heat Engine

French engineer, Sadi Carnot (1796-1832), in his study of the then-emerging heat engine technology crucial to the Industrial Revolution. He devised a theoretical cycle, now called the *Carnot cycle*, which is the most efficient cyclical process possible. The second law of thermodynamics can be restated in terms of the Carnot cycle, and so what Carnot actually discovered was this fundamental law. Any heat engine employing the Carnot cycle is called a *Carnot engine.*

The Second Law of Thermodynamics (third expression) *A Carnot engine operating between two given temperatures has the greatest possible efficiency of any heat engine operating between* *these two temperatures. Furthermore, all engines employing only reversible processes have this same maximum efficiency when operating between the same given temperatures.*

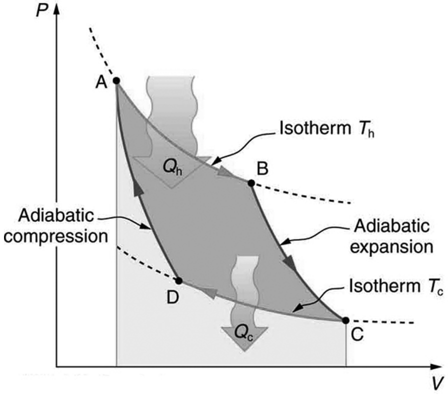

Figure 7.5 – *PV* diagram for a Carnot cycle, employing only reversible isothermal and adiabatic processes

Figure 7.5 shows the *PV* diagram for a Carnot cycle. The cycle comprises two isothermal and two adiabatic processes. Recall that both isothermal and adiabatic processes are, in principle, reversible.

Carnot also determined the efficiency of a perfect heat engine – that is, a Carnot engine. It is always true that the efficiency of a cyclical heat engine is given by (7.5).

What Carnot found was that for a perfect heat engine, the ratio Q_c/Q_h equals the ratio of the absolute temperatures of the heat reservoirs. That is, $Q_c/Q_h = T_c/T_h$ for a Carnot engine, so that the maximum or Carnot efficiency E_{ffC} is given by

$$E_{ffC} = 1 - \frac{T_c}{T_h}, \tag{7.6}$$

where T_h and T_c are in kelvins (or any other absolute temperature scale). No real heat engine can do as well as the Carnot efficiency – an actual efficiency of about 0.7 of this maximum is usually the best that can be accomplished.

7.3. Entropy

Recall that the simple definition of energy is the ability to do work. *Entropy* is a measure of how much energy is not available to do work. Although all forms of energy are interconvertible, and all can be used to do work, it is not always possible, even in principle, to convert the entire available energy into work. That is, unavailable energy is of interest in thermodynamics, because the field of thermodynamics arose from efforts to convert heat to work.

We noted that for a Carnot cycle, and hence for any reversible processes, $Q_c/Q_h = T_c/T_h$. Rearranging terms yields

$$\frac{Q_c}{T_c} = \frac{Q_h}{T_h}$$

for any reversible process. Q_c and Q_h are absolute values of the heat transfer at temperatures T_c and T_h, respectively. This ratio of Q/T is defined to be the change in entropy ΔS for a reversible process,

$$\Delta S = \left(\frac{Q}{T}\right)_{rev}, \tag{7.7}$$

where Q is the heat transfer which is positive for heat transfer into and negative for heat transfer out of, and T is the absolute temperature at which the reversible process takes place. The SI unit for entropy is joules per kelvin (J/K). If temperature changes during the process, then it is usually a good approximation (for small changes in temperature) to take T to be the average temperature, avoiding the need to use integral calculus to find ΔS:

$$\Delta S = \int_1^2 \frac{dQ}{T}. \tag{7.8}$$

The definition of ΔS is strictly valid only for reversible processes, such as used in a Carnot engine. However, we can find ΔS precisely even for real, irreversible processes. The reason is that the entropy S of a system, like internal energy U, depends only on the state of the system and not how it reached that condition. Entropy is a property of state. Thus the change in entropy ΔS of a system between state 1 and state 2 is the same no matter how the change occurs. The total change in entropy for a system in any reversible process is zero. There is an increase in entropy for any system undergoing an irreversible process.

The Second Law of Thermodynamics (fourth expression)
The total entropy of a system either increases or remains constant in any process; it never decreases.

For example, heat transfer cannot occur spontaneously from cold to hot, because entropy would decrease. Entropy is very different from energy. Entropy is not conserved but increases in all real processes. Reversible processes (such as in Carnot engines) are the processes in which the most heat transfer to work takes place and are also the ones that keep entropy constant.

When entropy increases, a certain amount of energy becomes permanently unavailable to do work. The energy is not lost, but its character is changed, so that some of it can never be converted to doing work – that is, to an organized force acting through a distance.

Order to Disorder

Entropy is related not only to the unavailability of energy to do work, it is also a measure of disorder. This notion was initially postulated by Ludwig Boltzmann in the 1800s. For example, melting a block of ice means taking a highly structured and orderly system of water molecules and converting it into a disorderly liquid in which molecules have no fixed positions. In another easily imagined example, suppose we mix equal masses of water originally at two different temperatures, say $20.0°C$ and $40.0°C$. The result is water at an intermediate temperature of $30.0°C$. Three outcomes have resulted: entropy has increased, some energy has become unavailable to do work, and the system has become less orderly. Let us think about each of these results. First, entropy has increased for the same reason that it did in the example above. Mixing the two bodies of water has the same effect as heat transfer from the hot one and the same heat transfer into the cold one. The mixing decreases the entropy of the hot water but increases the entropy of the cold water by a greater amount, producing an overall increase in entropy. Second, once the two masses of water are mixed, there is only one temperature – you cannot run a heat engine with them. The energy that could have been used to run a heat engine is now unavailable to do work. Third, the mixture is less orderly, or to use another term, less structured. Rather than having two masses at different temperatures and with different distributions of molecular speeds, we now have a single mass with a uniform temperature. These three results – entropy, unavailability of energy, and disorder – are not only related but are in fact essentially equivalent.

Statistical Interpretation of Entropy

A *macrostate* is an overall property of a system. It does not specify the details of the system. A *microstate* – a detailed description of every element of a system. Because counting microstates and macrostates involves statistics, this is called *statistical analysis.* The macrostates of a gas correspond to its macroscopic properties, such as volume, temperature, and pressure; and its microstates correspond to the detailed description of the positions and velocities of its atoms. Even a small amount of gas has a huge number of atoms: 1.0 cm^3 of an ideal gas at 1.0 atm and $0°C$ has 2.7×10^{19} atoms. So each macrostate has an immense number of microstates. In plain language, this means that there are an immense number of ways in which the atoms in a gas can be arranged, while still having the same pressure, temperature, and so on.

The most likely conditions (or macrostates) for a gas are those we see all the time – a random distribution of atoms in space with a Maxwell-Boltzmann distribution of speeds in random directions, as predicted by kinetic theory. This is the most disorderly and least structured condition we can imagine. In contrast, one type of very orderly and structured macrostate has all of the atoms in one corner of a container with identical velocities. There are very few ways to accomplish this (very few microstates corresponding to it), and so it is exceedingly unlikely ever to occur. (See Figure 7.6) Indeed, it is so unlikely that we have a law saying that it is impossible, which has never been observed to be violated – the second law of thermodynamics.

The disordered condition is one of high entropy, and the ordered one has low entropy. With a transfer of energy from another system, we could force all of the atoms into one corner and have a local decrease in entropy, but at the cost of an overall increase in entropy of the universe. If the atoms start out in one corner, they will quickly disperse and become uniformly distributed and will never return to the orderly original state (Figure 7.6(b)). Entropy will increase. With such a large sample of atoms, it is possible, but unimaginably unlikely, for entropy

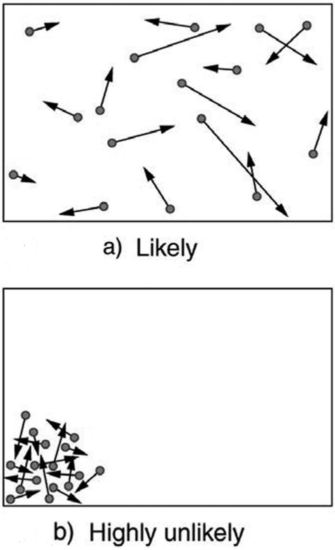

Figure 7.6 – a) The ordinary state of gas in a container is a disorderly, random distribution of atoms or molecules with a Maxwell-Boltzmann distribution of speeds. It is so unlikely that these atoms or molecules would ever end up in one corner of the container that it might as well be impossible. b) With energy transfer, the gas can be forced into one corner and its entropy greatly reduced. But left alone, it will spontaneously increase its entropy and return to the normal conditions, because they are immensely more likely

to decrease. Disorder is vastly more likely than order. The arguments that disorder and high entropy are the most probable states are quite convincing. The great Austrian physicist Ludwig Boltzmann (1844-1906), who, along with Maxwell, made so many contributions to kinetic theory, proved that the entropy of a system in a given state (a macrostate) can be written as

$$S = k \ln \Omega, \tag{7.9}$$

where k is Boltzmann's constant, and $\ln \Omega$ is the natural logarithm of the number of microstates Ω corresponding to the given macrostate. Ω is proportional to the probability that the macrostate will occur. Thus entropy is directly related to the probability of a state. The more likely the state, the greater its entropy. Boltzmann proved that this expression for S is equivalent to the definition $\Delta S = Q/T$, which we have used extensively.

Thus the second law of thermodynamics is explained on a very basic level: entropy either remains the same or increases in every process. This phenomenon is due to the extraordinarily small probability of a decrease, based on the extraordinarily larger number of microstates in systems with greater entropy. Entropy can decrease, but for any macroscopic system, this outcome is so unlikely that it will never be observed.

References

1. Fung, Y.C. Biomechanics. Mechanical Properties of Living Tissues. – New York : Springer-Verlag, 1993. – 568 p.

2. Glaser, R. Biophysics. An Introduction. – Berlin : Springer-Verlag, 2012. – 407 p.

3. College Physics // Textbook Equity Edition. – Houston : OpenStax College, 2013. – Vol. 1–3. – 1269 p.

Навчальне видання

Конспект лекцій із курсу
«Медична та біологічна фізика»
У двох частинах
Частина 1

Відповідальний за випуск І. В. Коплик
Редактор С. В. Чечоткіна
Комп'ютерне верстання А. М. Юнди

Формат 60×84/16. Ум. друк. арк. 7,91. Обл.-вид. арк. 7,29.

Видавець і виготовлювач
Сумський державний університет,
вул. Римського-Корсакова, 2, м. Суми, 40007
Свідоцтво суб'єкта видавничої справи ДК № 3062 від 17.12.2007.