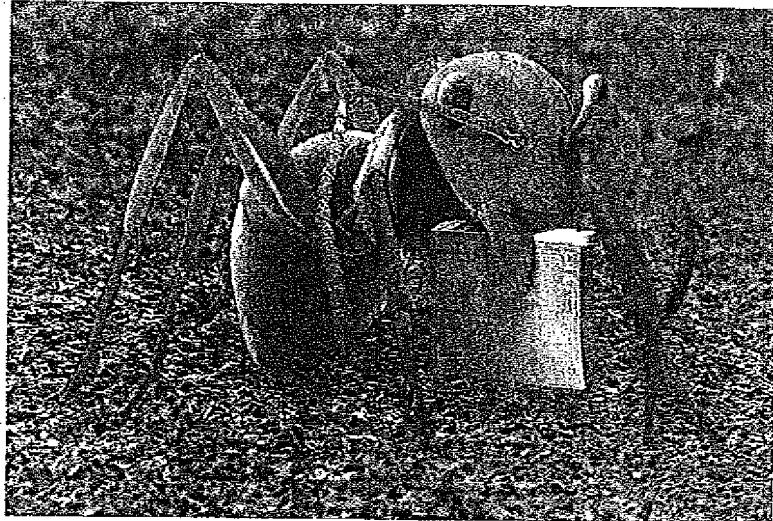


## CHAPTER 10

# The Solid State



Wood ant carrying a microchip that contains several million circuit elements.

### 10.1 CRYSTALLINE AND AMORPHOUS SOLIDS

*Long-range and short-range order*

### 10.2 IONIC CRYSTALS

*The attraction of opposites can produce a stable union*

### 10.3 COVALENT CRYSTALS

*Shared electrons lead to the strongest bonds*

### 10.4 VAN DER WAALS BOND

*Weak but everywhere*

### 10.5 METALLIC BOND

*A gas of free electrons is responsible for the characteristic properties of a metal*

### 10.6 BAND THEORY OF SOLIDS

*The energy band structure of a solid determines whether it is a conductor, an insulator, or a semiconductor*

### 10.7 SEMICONDUCTOR DEVICES

*The properties of the p-n junction are responsible for the microelectronics industry*

### 10.8 ENERGY BANDS: ALTERNATIVE ANALYSIS

*How the periodicity of a crystal lattice leads to allowed and forbidden bands*

### 10.9 SUPERCONDUCTIVITY

*No resistance at all, but only at very low temperatures (so far)*

### 10.10 BOUND ELECTRON PAIRS

*The key to superconductivity*

A solid consists of atoms, ions, or molecules packed closely together, and the forces that hold them in place give rise to the distinctive properties of the various kinds of solid. The covalent bonds that can link a fixed number of atoms to form a certain molecule can also link an unlimited number of them to form a solid. In addition, ionic, van der Waals, and metallic bonds provide the cohesive forces in solids whose structural elements are respectively ions, molecules, and metal atoms. All these bonds involve electric forces, with the chief differences among them being in the ways in which the outer electrons of the structural elements are distributed. Although very little of the matter in the universe is in the solid state, solids constitute much of the physical world around us, and a large part of modern technology is based on the special characteristics of various solid materials.

## 10.1 CRYSTALLINE AND AMORPHOUS SOLIDS

### *Long-range and short-range order*

Most solids are crystalline, with the atoms, ions, or molecules of which they are composed falling into regular, repeated three-dimensional patterns. The presence of long-range order is thus the defining property of a crystal, although relatively few samples of crystalline solids consist of single crystals. Most are polycrystalline and are composed of a great many small crystals (sometimes called crystallites).

Other solids lack the definite arrangements of their member particles so conspicuous in crystals. They may be regarded as supercooled liquids whose stiffness is due to an exceptionally high viscosity. Glass, pitch, and many plastics are examples of such amorphous ("without form") solids.

Amorphous solids do exhibit short-range order in their structures, however. The distinction between the two kinds of order is nicely exhibited in boron trioxide ( $B_2O_3$ ), which can occur in both crystalline and amorphous forms. In each case every boron atom is surrounded by three oxygen atoms, which represents a short-range order. In a  $B_2O_3$  crystal a long-range order is also present, as shown in a two-dimensional representation in Fig. 10.1. Amorphous  $B_2O_3$ , a vitreous or "glassy" substance, lacks this additional regularity. Crystallization from the vitreous state is so sluggish that it ordinarily does not occur, but it is not unknown. Glass may devitrify when heated until it has not quite begun to soften, and extremely old glass specimens are sometimes found to have crystallized.

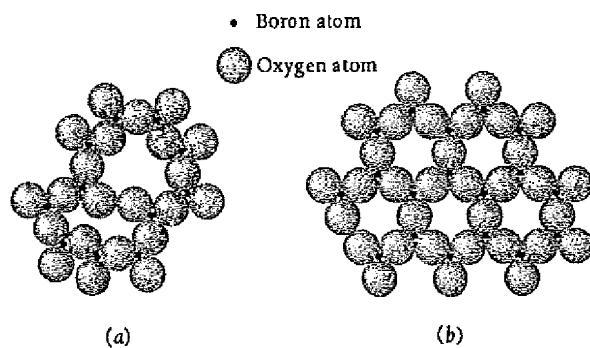


Figure 10.1 Two-dimensional representation of  $B_2O_3$ . (a) Amorphous  $B_2O_3$  exhibits only short-range order. (b) Crystalline  $B_2O_3$  exhibits long-range order as well.

The analogy between an amorphous solid and a liquid helps in understanding both states of matter. The density of a given liquid is usually close to that of the corresponding solid, for instance, which suggests that the degree of packing is similar. This inference is supported by the compressibilities of these states. Furthermore, x-ray diffraction indicates that many liquids have definite short-range structures at any instant, quite similar to those of amorphous solids except that the groupings of liquid molecules are continually shifting. A conspicuous example of short-range order in a liquid occurs in water just above the melting point, where the result is a lower density than at higher temperatures because  $H_2O$  molecules are less tightly packed when linked in crystals than when free to move.

The bonds in an amorphous solid vary in strength because of the lack of long-range order. When an amorphous solid is heated, the weakest bonds break at lower temperatures than the others, and the solid softens gradually. In a crystalline solid the bonds break simultaneously, and melting has a sudden onset. Metallic "glasses" have been made from mixtures of metals whose atoms differ greatly in size, which prevents them from forming the ordered structures of crystals when cooled from a molten state. One such metallic glass has half the density of steel but twice its strength, and is hard but can be deformed without breaking. Its gradual softening when heated make the material exceptionally easy to shape.

### Crystal Defects

In a perfect crystal each atom has a definite equilibrium location in a regular array. Actual crystals are never perfect. Defects such as missing atoms, atoms out of place, irregularities in the spacing of rows of atoms and the presence of impurities have a considerable bearing on the physical properties of a crystal. Thus the behavior of a solid under stress is largely determined by the nature and concentration of defects in its structure, as is the electrical behavior of a semiconductor.

The simplest category of crystal imperfection is the point defect. Figure 10.2 shows the basic kinds of point defect. Both vacancies and interstitials, which require about 1 to 2 eV to be created, occur in all crystals as a result of thermal excitation, and their number accordingly increases rapidly with temperature. Of much importance is the production of such defects by particle radiation. In a nuclear reactor, for instance, energetic

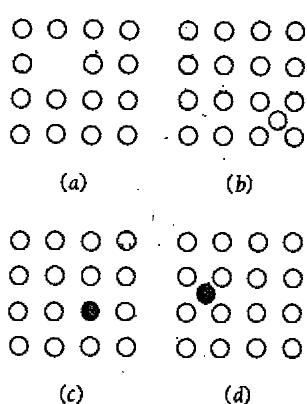
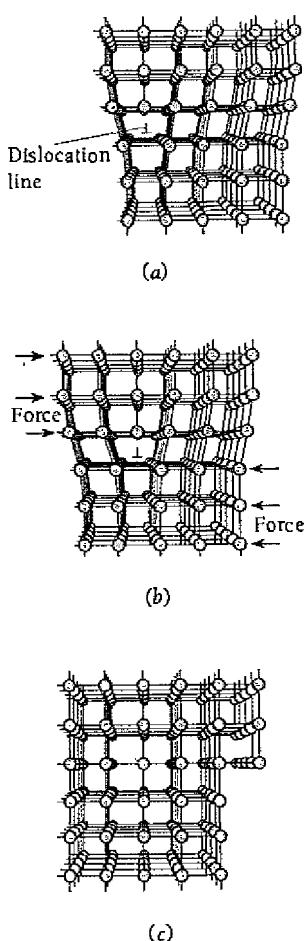


Figure 10.2 Point defects in a crystal. (a) Vacancy. (b) Interstitial. (c) Substitutional impurity. (d) Interstitial impurity.



**Figure 10.3** A crystal under stress becomes permanently deformed when dislocations in its structure shift their positions. (a) Initial configuration of a crystal with an edge dislocation. (b) The dislocation moves to the right as the atoms in the layer under it successively shift their bonds with those of the upper layer one line at a time. (c) The crystal has taken on a permanent deformation. The forces needed for this step-by-step process are much smaller than those needed to slide one entire layer of atoms past another layer.

neutrons readily knock atoms out of their normal locations. The result is a change in the properties of the bombarded material; most metals, for instance, become more brittle.

The effects of impurity atoms on the electrical properties of semiconductors, which underlie the operation of such devices as transistors, are discussed later in this chapter.

A dislocation is a type of crystal defect in which a line of atoms is not in its proper position. Dislocations are of two basic kinds. Figure 10.3 shows an edge dislocation, which we can visualize as the result of removing part of a layer (here vertical) of atoms. Edge dislocations enable a solid to be permanently deformed without breaking, a property called ductility. Metals are the most ductile solids. In the figure the bonds between atoms are represented by lines. The other kind of dislocation is the screw dislocation. We can visualize the formation of a screw dislocation by imagining that a cut is made partway into a perfect crystal and one side of the cut is then displaced relative to the other, as in Fig. 10.4. The atomic layers spiral around the dislocation, which accounts for its name. Actual dislocations in crystals are usually combinations of the edge and screw varieties.

Dislocations multiply when a solid is deformed. When the dislocations become so numerous and tangled together that they impede one another's motion, the material is then less easy to deform. This effect is called work hardening. Strongly heating (annealing) a work-hardened solid tends to return its disordered lattice to regularity and it becomes more ductile as a result. Steel bars and sheets formed by cold rolling are much harder than those formed by hot rolling.

## 10.2 IONIC CRYSTALS

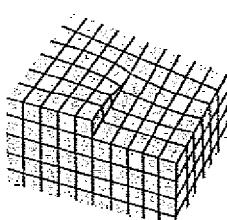
*The attraction of opposites can produce a stable union*

Ionic bonds come into being when atoms that have low ionization energies, and hence lose electrons readily, interact with other atoms that tend to acquire excess electrons. The former atoms give up electrons to the latter, and they thereupon become positive and negative ions respectively (Fig. 8.2). In an ionic crystal these ions assemble themselves in an equilibrium configuration in which the attractive forces between positive and negative ions balance the repulsive forces between the ions.

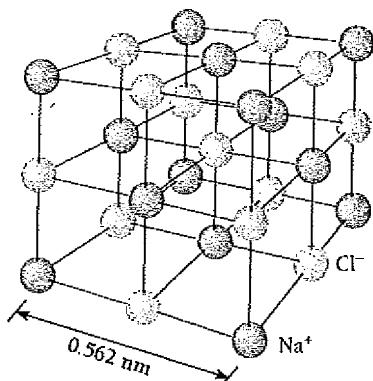
As in the case of molecules, crystals of all types are prevented from collapsing under the influence of the cohesive forces present by the action of the exclusion principle, which requires the occupancy of higher energy states when electron shells of different atoms overlap and mesh together.

In general, in an ionic crystal each ion is surrounded by as many ions of the opposite sign as can fit closely, which leads to maximum stability. The relative sizes of the ions involved therefore govern the type of structure that occurs. Two common types of structure found in ionic crystals are shown in Figs. 10.5 and 10.6.

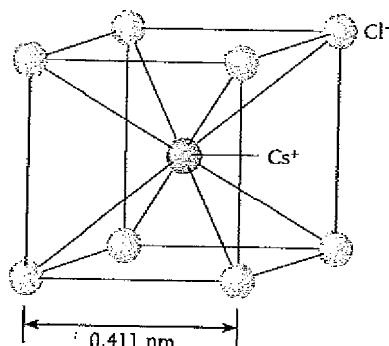
Ionic bonds between the atoms of two elements can form when one element has a low ionization energy, so that its atoms tend to become positive ions, and the other



**Figure 10.4** A screw dislocation.



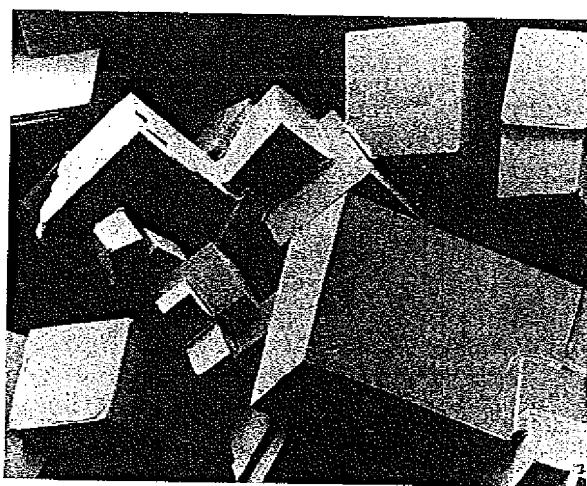
**Figure 10.5** The face-centered cubic structure of NaCl. The coordination number (the number of nearest neighbors about each ion) is 6.



**Figure 10.6** The body-centered cubic structure of CsCl. The coordination number is 8.

element has a high electron affinity. Electron affinity is the energy released when an electron is added to an atom of a given element; the greater the electron affinity, the more such atoms tend to become negative ions. Sodium, with an ionization energy of 5.14 eV, tends to form  $\text{Na}^+$  ions, and chlorine, with an electron affinity of 3.61 eV, tends to form  $\text{Cl}^-$  ions. The condition for a stable crystal of NaCl is simply that the total energy of a system of  $\text{Na}^+$  and  $\text{Cl}^-$  ions be less than the total energy of a system of Na and Cl atoms.

The cohesive energy of an ionic crystal is the energy per ion needed to break the crystal up into individual atoms. Part of the cohesive energy is the electric potential energy  $U_{\text{coulomb}}$  of the ions. Let us consider a  $\text{Na}^+$  ion in NaCl. From Fig. 10.5 its nearest neighbors are six  $\text{Cl}^-$  ions, each one the distance  $r$  away. The potential energy



Electron micrograph of sodium chloride crystals. The cubic structure of the crystals is often disrupted by dislocations.

of the  $\text{Na}^+$  ion due to these six  $\text{Cl}^-$  ions is therefore

$$U_1 = -\frac{6e^2}{4\pi\epsilon_0 r}$$

The next nearest neighbors are 12  $\text{Na}^+$  ions, each one the distance  $\sqrt{2} r$  away since the diagonal of a square  $r$  long on a side is  $\sqrt{2} r$ . The potential energy of the  $\text{Na}^+$  ion due to the 12  $\text{Na}^+$  ions is

$$U_2 = +\frac{12e^2}{4\pi\epsilon_0 \sqrt{2} r}$$

When the summation is continued over all the + and - ions in a crystal of infinite size, the result is

$$U_{\text{coulomb}} = -\frac{e^2}{4\pi\epsilon_0 r} \left( 6 - \frac{12}{\sqrt{2}} + \dots \right) = -1.748 \frac{e^2}{4\pi\epsilon_0 r}$$

or, in general,

$$\text{Coulomb energy} \quad U_{\text{coulomb}} = -\alpha \frac{e^2}{4\pi\epsilon_0 r} \quad (10.1)$$

This result holds for the potential energy of a  $\text{Cl}^-$  ion as well, of course.

The quantity  $\alpha$  is called the **Madelung constant** of the crystal, and it has the same value for all crystals of the same structure. Similar calculations for other crystal varieties yield different Madelung constants. Crystals whose structures are like that of cesium chloride (Fig. 10.6), for instance, have  $\alpha = 1.763$ . Simple crystal structures have Madelung constants that lie between 1.6 and 1.8.

The potential energy contribution of the repulsive forces due to the action of the exclusion principle has the approximate form

$$\text{Repulsive energy} \quad U_{\text{repulsive}} = \frac{B}{r^n} \quad (10.2)$$

The sign of  $U_{\text{repulsive}}$  is positive, which corresponds to a repulsion. The dependence on  $r^{-n}$  implies a short-range force that increases as the interionic distance  $r$  decreases. The total potential energy of each ion due to its interactions with all the other ions is therefore

$$U_{\text{total}} = U_{\text{coulomb}} + U_{\text{repulsive}} = -\frac{\alpha e^2}{4\pi\epsilon_0 r} + \frac{B}{r^n} \quad (10.3)$$

How can we find the value of  $B$ ? At the equilibrium separation  $r_0$  of the ions,  $U$  is a minimum by definition, and so  $dU/dr = 0$  when  $r = r_0$ . Hence

$$\left( \frac{dU}{dr} \right)_{r=r_0} = \frac{\alpha e^2}{4\pi\epsilon_0 r_0^2} - \frac{nB}{r_0^{n+1}} = 0$$

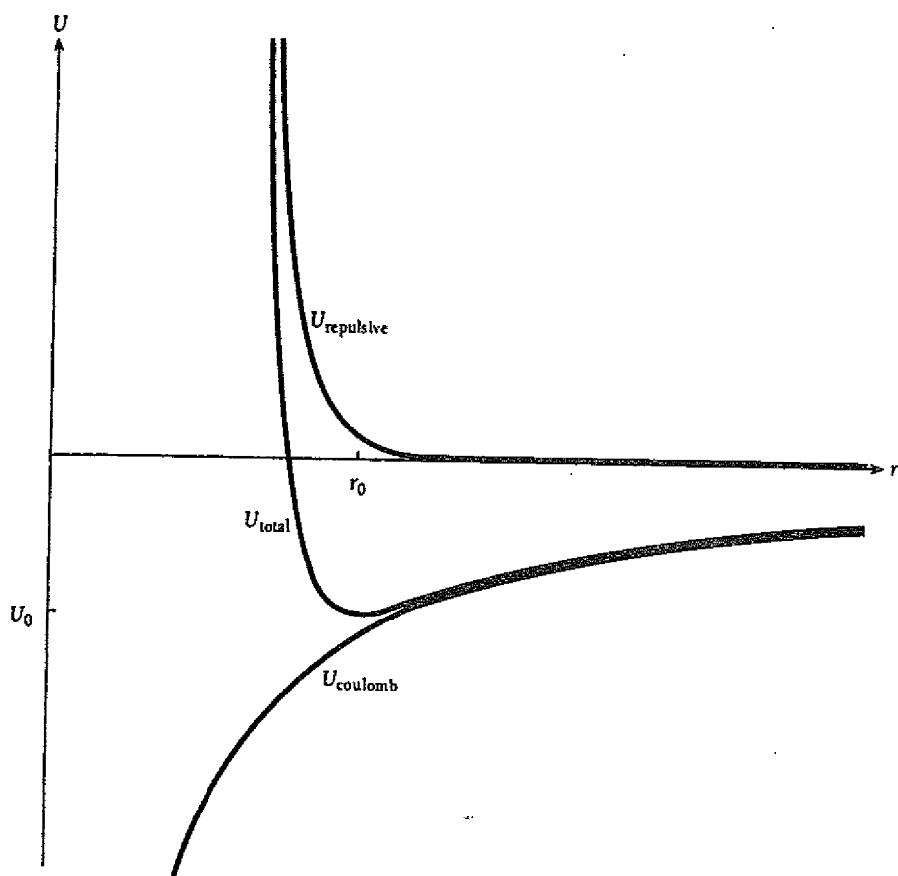


Figure 10.7 How the ionic potential energies in an ionic crystal vary with ionic separation  $r$ . The minimum value  $U_0$  of  $U_{\text{total}}$  occurs at an equilibrium separation of  $r_0$ .

$$B = \frac{\alpha e^2}{4\pi\epsilon_0 n} r_0^{n-1} \quad (10.4)$$

The total potential energy at the equilibrium separation is therefore given by

$$\begin{aligned} &\text{Total} \\ &\text{potential} \\ &\text{energy} \end{aligned} \quad U_0 = -\frac{\alpha e^2}{4\pi\epsilon_0 r_0} \left(1 - \frac{1}{n}\right) \quad (10.5)$$

We must add this amount of energy per ion pair to separate an ionic crystal into individual ions (Fig. 10.7). For the cohesive energy, which corresponds to separating the crystal into atoms, we must take into account the energy involved in shifting an electron from a Na atom to a Cl atom to give a  $\text{Na}^+ \cdot \text{Cl}^-$  ion pair.

The exponent  $n$  can be found from the observed compressibilities of ionic crystals. The average result is  $n \approx 9$ , which means that the repulsive force varies sharply with  $r$ . The ions are "hard" rather than "soft" and strongly resist being packed too tightly. At the equilibrium ion spacing, the mutual repulsion due to the exclusion principle (as

distinct from the electric repulsion between like ions) decreases the potential energy by about 11 percent. A really precise knowledge of  $n$  is not essential; if  $n = 10$  instead of  $n = 9$ ,  $U_0$  would change by only 1 percent.

### Example 10.1

In an NaCl crystal, the equilibrium distance  $r_0$  between ions is 0.281 nm. Find the cohesive energy in NaCl.

#### Solution

Since  $\alpha = 1.748$  and  $n \approx 9$ , the potential energy per ion pair is

$$U_0 = -\frac{\alpha e^2}{4\pi\epsilon_0 r_0} \left(1 - \frac{1}{n}\right) = -\frac{(9 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2)(1.748)(1.60 \times 10^{-19} \text{ C})^2}{2.81 \times 10^{-10} \text{ m}} \left(1 - \frac{1}{9}\right)$$

$$= -1.27 \times 10^{-18} \text{ J} = -7.96 \text{ eV}$$

Half this figure,  $-3.98 \text{ eV}$ , represents the contribution per ion to the cohesive energy of the crystal.

Now we need the electron transfer energy, which is the sum of the  $+5.14\text{-eV}$  ionization energy of Na and the  $-3.61\text{-eV}$  electron affinity of Cl, or  $+1.53 \text{ eV}$ . Each atom therefore contributes  $+0.77 \text{ eV}$  to the cohesive energy from this source. The total cohesive energy per atom is thus

$$E_{\text{cohesive}} = (-3.98 + 0.77) \text{ eV} = -3.21 \text{ eV}$$

which is not far from the experimental value of  $-3.28 \text{ eV}$ .

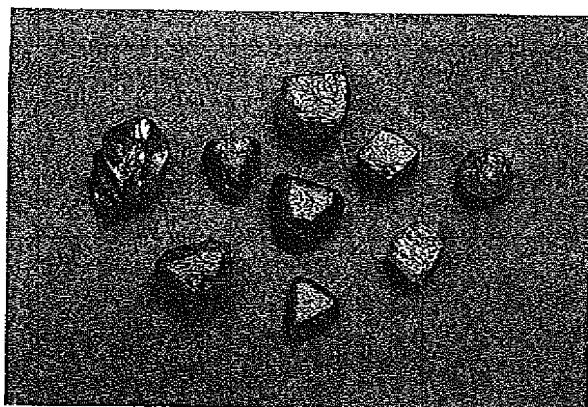
Most ionic solids are hard, owing to the strength of the bonds between their constituent ions, and have high melting points. They are usually brittle as well, since the slipping of atoms past one another that accounts for the ductility of metals is prevented by the ordering of positive and negative ions imposed by the nature of the bonds. Polar liquids such as water are able to dissolve many ionic crystals, but covalent liquids such as gasoline generally cannot. Because the outer electrons of their ions are tightly bound, ionic crystals are good electrical insulators and are transparent to visible light. However, such crystals strongly absorb infrared radiation at the frequencies at which the ions vibrate about their equilibrium positions.

### 10.3 COVALENT CRYSTALS

#### *Shared electrons lead to the strongest bonds*

The cohesive forces in covalent crystals arise from the sharing of electrons by adjacent atoms. Each atom that participates in a covalent bond contributes an electron to the bond. Figure 10.8 shows the tetrahedral structure of a diamond crystal, each of whose carbon atoms is linked by covalent bonds to four other carbon atoms.

Another crystalline form of carbon is graphite. Graphite consists of layers of carbon atoms in a hexagonal network in which each atom is joined to three others by covalent bonds  $120^\circ$  apart, as in Fig. 10.9. One electron per atom participates in



Uncut diamonds. The strength of the covalent bonds between adjacent carbon atoms gives diamonds their hardness.

each bond. This leaves one outer electron in each carbon atom free to circulate through the network, thereby accounting for graphite's near-metallic luster and electrical conductivity. Although each layer is quite strong, weak van der Waals forces (Sec. 10.4) bond the layers together. As a result the layers can slide past each other readily and are easily flaked apart, which is why graphite is so useful as a lubricant and in pencils.

Under ordinary conditions graphite is more stable than diamond, so crystallizing carbon normally produces only graphite. Because graphite is less dense than diamond ( $2.25 \text{ g/cm}^3$  versus  $3.51 \text{ g/cm}^3$ ), high pressures favor the formation of diamond. Natural diamonds originated deep in the earth where pressures are enormous. To synthesize diamonds, graphite is dissolved in molten cobalt or nickel and the mixture is compressed at about  $1600 \text{ K}$  to about 60,000 bar. The resulting diamonds are less than 1 mm across and are widely used industrially for cutting and grinding tools.

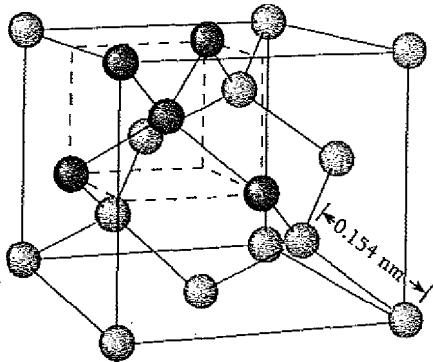


Figure 10.8 The tetrahedral structure of diamond. The coordination number is 4.

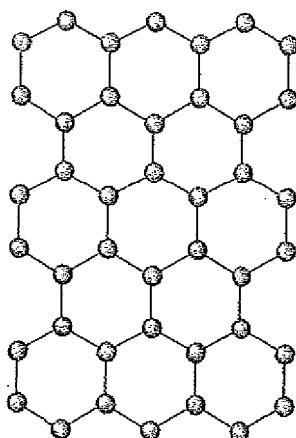


Figure 10.9 Graphite consists of layers of carbon atoms in hexagonal arrays, with each atom bonded to three others. The layers are held together by weak van der Waals forces.

Purely covalent crystals are relatively few in number. In addition to diamond, some examples are silicon, germanium, and silicon carbide, all of which have the same tetrahedral structure as diamond; in SiC each atom is surrounded by four atoms of the other kind. Cohesive energies are usually greater in covalent crystals than in ionic ones. As a result covalent crystals are hard (diamond is the hardest substance known, and SiC is the industrial abrasive carborundum), have high melting points, and are insoluble in all ordinary liquids. The optical and electrical properties of covalent solids are discussed later.

### Buckyballs and Nanotubes

**A**n unexpected form of carbon was accidentally discovered in 1985 at Rice University in Texas. The commonest version consists of 60 carbon atoms arranged in a cage structure of 12 pentagons and 20 hexagons whose geometry is like that of a soccer ball (Fig. 10.10). This extraordinary molecule was called "buckminsterfullerene" in honor of the American architect R. Buckminster Fuller, whose geodesic domes it resembles; the name is usually shortened to buckyball.

Buckyballs, which are stable and chemically unreactive, can be made in the laboratory from graphite and are present in small quantities in ordinary soot and in a carbon-rich rock found in Russia. The original  $C_{60}$  buckyball is not the only form of fullerene known:  $C_{28}$ ,  $C_{32}$ ,  $C_{50}$ ,  $C_{70}$ , and still larger ones have been made. Fullerene molecules are held together to form solids by van der Waals bonds like those that hold together the layers of C atoms in graphite. Since their discovery, the fullerenes and their offshoots have shown some remarkable properties. For instance, the combination of  $C_{60}$  with potassium to form  $K_3C_{60}$  yields a superconductor at low temperatures.

Carbon nanotubes, cousins of buckyballs, consist of tiny cylinders of carbon atoms arranged in hexagons, like rolled-up chicken wire. Depending on whether their rows of hexagons are straight or wind around in a helix, such nanotubes act either as electrical conductors or as semiconductors and their use is being explored in such electronic applications as transistors and flat-panel displays. If carbon nanotubes can be made long enough, they will form exceedingly strong fibers, ten times stronger than steel while six times lighter, that are flexible as well. Fibers like this would be ideal in composite materials to reinforce epoxy resins. Nanotubes also have promise for storing the hydrogen needed for the fuel cells of future electric cars, which would make heavy steel containers unnecessary.

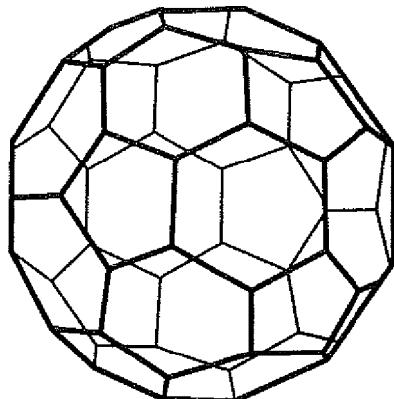


Figure 10.10 In a buckyball, carbon atoms form a closed cagelike structure in which each atom is bonded to three others. Shown here is the  $C_{60}$  buckyball that contains 60 carbon atoms. The lines represent carbon-carbon bonds; their pattern of hexagons and pentagons closely resembles the pattern made by the seams of a soccer ball. Other buckyballs have different numbers of carbon atoms.

## 10.4 VAN DER WAALS BOND

*Weak but everywhere*

All atoms and molecules—even inert-gas atoms such as those of helium and argon—exhibit weak, short-range attractions for one another due to van der Waals forces. These forces were proposed over a century ago by the Dutch physicist Johannes van der Waals to explain departures of real gases from the ideal-gas law. The explanation of the actual mechanism of the forces, of course, is more recent.

Van der Waals forces are responsible for the condensation of gases into liquids and the freezing of liquids into solids in the absence of ionic, covalent, or metallic bonding mechanisms. Such familiar aspects of the behavior of matter in bulk as friction, surface tension, viscosity, adhesion, cohesion, and so on, also arise from these forces. As we shall find, the van der Waals attraction between two molecules the distance  $r$  apart is proportional to  $r^{-7}$ , so that it is significant only for molecules very close together.

We begin by noting that many molecules, called **polar molecules**, have permanent electric dipole moments. An example is the  $\text{H}_2\text{O}$  molecule, in which the concentration of electrons around the oxygen atom makes that end of the molecule more negative than the end where the hydrogen atoms are. Such molecules tend to clump together with ends of opposite sign adjacent, as in Fig. 10.11.

A polar molecule can also attract molecules which lack a permanent dipole moment. The process is illustrated in Fig. 10.12. The electric field of the polar molecule causes a separation of charge in the other molecule, with the induced moment the same in direction as that of the polar molecule. The result is an attractive force. The effect is the same as that involved in the attraction of an unmagnetized piece of iron by a magnet.

Let us see what the characteristics of the attractive force between a polar and a non-polar molecule depend on. The electric field  $E$  a distance  $r$  from a dipole of moment  $p$  is given by

$$\text{Dipole electric field} \quad E = \frac{1}{4\pi\epsilon_0} \left[ \frac{p}{r^3} - \frac{3(p \cdot r)}{r^5} r \right] \quad (10.6)$$

We recall from vector analysis that  $p \cdot r = pr \cos\theta$ , where  $\theta$  is the angle between  $p$  and  $r$ . The field  $E$  induces in the other, normally nonpolar molecule an electric dipole moment  $p'$  proportional to  $E$  in magnitude and ideally in the same direction. Hence

$$\text{Induced dipole moment} \quad p' = \alpha E \quad (10.7)$$

where  $\alpha$  is a constant called the **polarizability** of the molecule. The energy of the induced dipole in the electric field  $E$  is

$$\begin{aligned} U &= -p' \cdot E = -\alpha E \cdot E \\ &= -\frac{\alpha}{(4\pi\epsilon_0)^2} \left( \frac{p^2}{r^6} - \frac{3p^2}{r^6} \cos^2\theta - \frac{3p^2}{r^6} \cos^2\theta + \frac{9p^2}{r^6} \cos^2\theta \right) \\ \text{Interaction energy} &= -\frac{\alpha}{(4\pi\epsilon_0)^2} (1 + 3 \cos^2\theta) \frac{p^2}{r^6} \end{aligned} \quad (10.8)$$

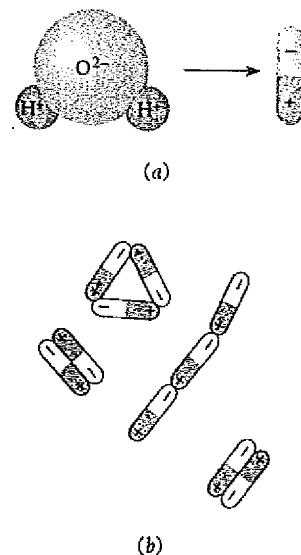


Figure 10.11 (a) The water molecule is polar because the end where the H atoms are attached behaves as if positively charged and the opposite end behaves as if negatively charged. (b) Polar molecules attract each other.

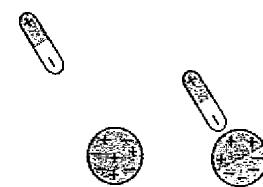


Figure 10.12 Polar molecules attract polarizable molecules.

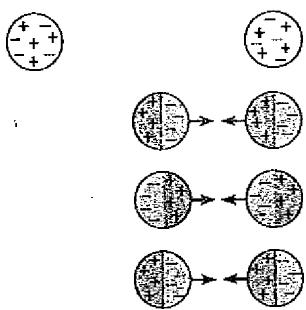


Figure 10.13 On the average, nonpolar molecules have symmetrical charge distributions, but at any moment the distributions are asymmetric. The fluctuations in the charge distributions of nearby molecules are coordinated as shown. This situation leads to an attractive force between them whose magnitude varies as  $1/r^7$ , where  $r$  is their distance apart.

The potential energy of the two molecules that arises from their interaction is negative, signifying that the force between them is attractive, and is proportional to  $r^{-6}$ . The force itself is equal to  $-dU/dr$  and so is proportional to  $r^{-7}$ , which means that it drops rapidly with increasing separation. Doubling the distance between two molecules reduces the attractive force between them to only 0.8 percent of its original value.

More remarkably, two nonpolar molecules can attract each other by the above mechanism. The electron distribution in a nonpolar molecule is symmetric on the average. However, the electrons themselves are in constant motion and at any given instant one part or another of the molecule has an excess of them. Instead of the fixed charge asymmetry of a polar molecule, a nonpolar molecule has a constantly shifting asymmetry. When two nonpolar molecules are close enough, their fluctuating charge distributions tend to shift together with adjacent ends always having opposite sign (Fig. 10.13), which leads to an attractive force.

Van der Waals forces occur not only between all molecules but also between all atoms, including those of the rare gases which do not otherwise interact. Without such forces these gases would not condense into liquids or solids. The values of  $p^2$  (or  $\bar{p}^2$ , the average of  $p^2$ , which applies for molecules with no permanent dipole moment) and the polarizability  $\alpha$  are comparable for most molecules. This is part of the reason why the densities and heats of vaporization of liquids, properties that depend on the strength of intermolecular forces, have a rather narrow range.

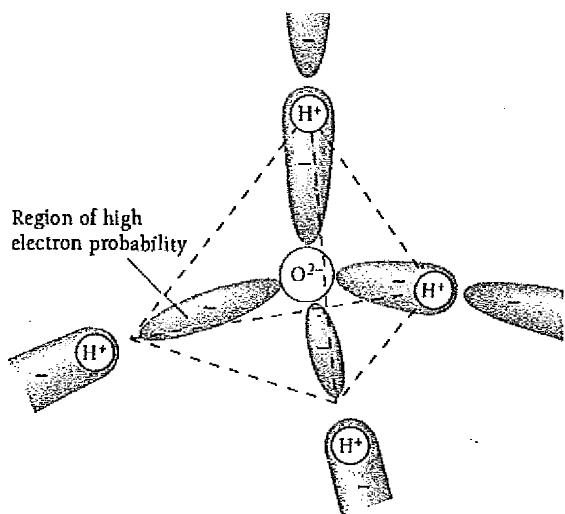
Van der Waals forces are much weaker than those found in ionic and covalent bonds, and as a result molecular crystals generally have low melting and boiling points and little mechanical strength. Cohesive energies are low, only 0.08 eV/atom in solid argon (melting point  $-189^\circ\text{C}$ ), 0.01 eV/molecule in solid hydrogen (mp  $-259^\circ\text{C}$ ), and 0.1 eV/molecule in solid methane,  $\text{CH}_4$  (mp  $-183^\circ\text{C}$ ).

### Hydrogen Bonds

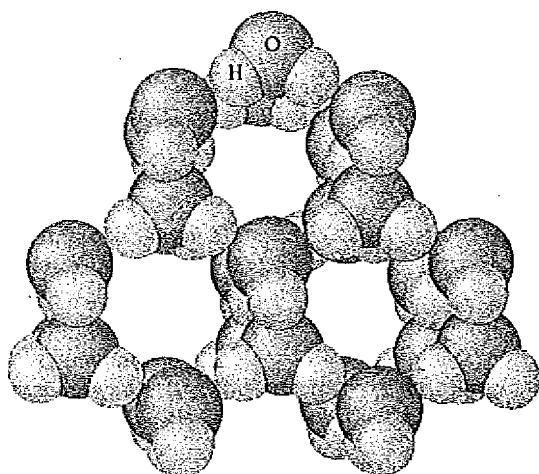
An especially strong type of van der Waals bond called a hydrogen bond occurs between certain molecules containing hydrogen atoms. The electron distribution in such a molecule is severely distorted by the affinity of a heavier atom for electrons. Each hydrogen atom in effect donates most of its negative charge to the other atom, to leave behind a poorly shielded proton. The result is a molecule with a localized positive charge which can link up with the concentration of negative charge elsewhere in another molecule of the same kind. The key factor here is the small effective size of the poorly shielded proton, since electric forces vary as  $1/r^2$ . Hydrogen bonds are typically about a tenth as strong as covalent bonds.

Water molecules are exceptionally prone to form hydrogen bonds because the electrons around the O atom in  $\text{H}_2\text{O}$  are not symmetrically distributed but are more likely to be found in certain regions of high probability density. These regions project outward as though toward the vertices of a tetrahedron, as shown in Fig. 10.14. Hydrogen atoms are at two of these vertices, which accordingly exhibit localized positive charges, while the other two vertices exhibit somewhat more diffuse negative charges.

Each  $\text{H}_2\text{O}$  molecule can therefore form hydrogen bonds with four other  $\text{H}_2\text{O}$  molecules. In two of these bonds the central molecule provides the bridging protons, and in the other two the attached molecules provide them. In the liquid state, the hydrogen bonds between adjacent  $\text{H}_2\text{O}$  molecules are continually being broken and re-formed owing to thermal agitation, but even so at any instant the molecules are combined in definite clusters. In the solid state, these clusters are large and stable and



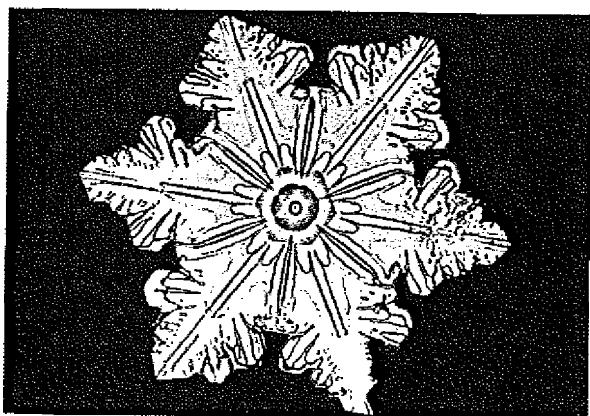
**Figure 10.14** In an  $\text{H}_2\text{O}$  molecule, the four pairs of valence electrons around the oxygen atom (six electrons contributed by the O atom and one each by the H atoms) preferentially occupy four regions that form a tetrahedral pattern. Each  $\text{H}_2\text{O}$  molecule can form hydrogen bonds with four other  $\text{H}_2\text{O}$  molecules.



**Figure 10.15** The structure of an ice crystal, showing the open hexagonal arrangement of the  $\text{H}_2\text{O}$  molecules. There is less order in liquid water, which allows the molecules to be closer together on the average than they are in ice. Thus the density of ice is less than that of water, and ice floats.

constitute ice crystals (Fig. 10.15). With only four nearest neighbors around each molecule; instead of as many as twelve in other solids, ice crystals have extremely open structures, which is why ice has a relatively low density.

Hydrogen bonds occur widely in biological materials. The peptide bonds that join amino acids to form proteins are hydrogen bonds, for example, as are the bonds that hold together the two strands of the double helix of DNA. The bonds in DNA are strong enough for it to be a reliable store of genetic information but weak enough to permit its strands to be unzipped temporarily for the information to be transcribed ultimately into proteins and also permanently for DNA replication.



The water molecules in a snowflake are held together by hydrogen bonds.

## 10.5 METALLIC BOND

*A gas of free electrons is responsible for the characteristic properties of a metal*

The valence (outer) electrons of metal atoms are only weakly bound, as Fig. 7.10 shows. When such atoms interact to become a solid, their valence electrons form a "gas" of electrons that move with relative freedom through the resulting assembly of metal ions. The electron gas acts to hold the ions together and also provides the high electric and thermal conductivities, opacity, surface luster, and other characteristic properties of metals. Because the free electrons do not belong to particular atom-atom bonds, different metals can be alloyed together in more-or-less arbitrary proportions if their atoms are similar in size. In contrast, the components of ionic solids and of covalent solids such as SiC combine only in specific proportions.

As in any other solid, metal atoms cohere because their total energy is lower when they are bound together than when they are separate atoms. This energy reduction occurs in a metal crystal because each valence electron is on the average closer to one ion or another than it would be if it belonged to an isolated atom. Hence the electron's potential energy is less in the crystal than in the atom.

Another factor is involved here: although the potential energy of the free electrons is reduced in a metal crystal, their kinetic energy is increased. The valence energy levels of the metal atoms are all slightly altered by their interactions to give as many different energy levels as the total number of atoms present. The levels are so closely spaced as to form an essentially continuous energy band. As discussed in Chap. 9, the free electrons in this band have a Fermi-Dirac energy distribution in which, at 0 K, their kinetic energies range from 0 to a maximum of  $\epsilon_F$ , the Fermi energy. The Fermi energy in copper, for example, is 9.04 eV, and the average KE of the free electrons in metallic copper at 0 K is 4.22 eV.

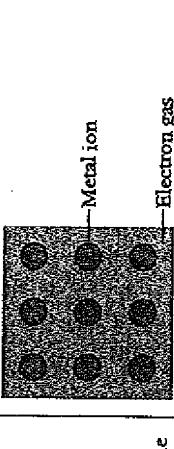
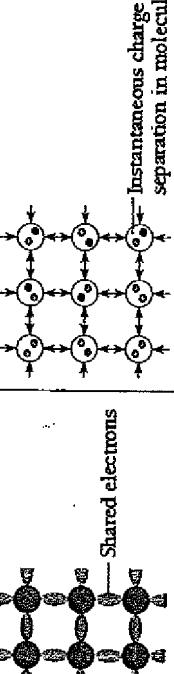
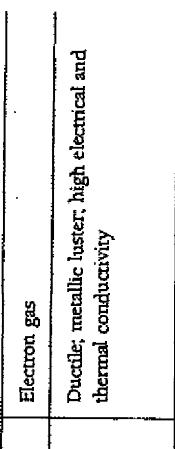
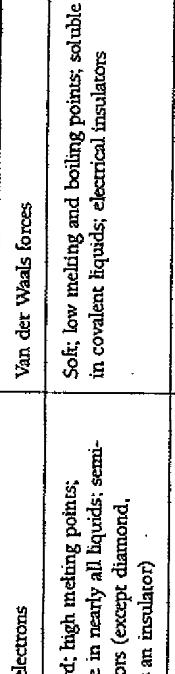
### Metallic Hydrogen

**H**ydrogen is in group 1 of the periodic table, all the other elements of which are metals. Hydrogen is an exception, which is not surprising when it is in the gaseous state, but it does not behave as a metal (for instance by being a good electrical conductor) even when it has been cooled to the liquid or solid states. The reason is that both liquid and solid hydrogen at atmospheric pressure consist of hydrogen molecules,  $H_2$ , and these molecules hold their electrons so tightly that none can break loose and move about freely as in the case of the atomic electrons of metals.

However, extremely high pressures—several million times atmospheric pressure—turn hydrogen into a conducting liquid. What the pressure does is force the  $H_2$  molecules so close together that their electron wave functions overlap, which allows electrons to migrate from one molecule to the next. Pressures inside the giant planet Jupiter, which consists largely of hydrogen, are sufficient for Jupiter apparently to have a hydrogen core that is in the form of a liquid metal. Electric currents in Jupiter's core produce its magnetic field; this field is about 20 times stronger than the earth's field, which is due to currents in its molten iron core.

Conceivably someday solid metallic hydrogen could be created, perhaps combined with other substances to help stabilize it, that would survive ordinary temperatures and pressures. The possible properties of such metallic hydrogen include superconductivity and light weight combined with mechanical strength. The energy that would be released by allowing solid hydrogen to turn into a gas could be used to propel spacecraft—it might give five times as much thrust per kilogram as current rocket fuels. Because solid hydrogen would be much denser than ordinary hydrogen, in the form of its isotopes deuterium and tritium it would make an extremely efficient fuel for fusion reactors. All in all, wonderful prospects, but how realistic they are remains to be seen.

**Table 10.1** Types of Crystalline Solids. The cohesive energy is the work needed to remove an atom (or molecule) from the crystal and so indicates the strength of the bonds holding it in place.

Type	Ionic	Covalent	Molecular	Metallic
Lattice				
Bond	Electric attraction.	Shared electrons	Van der Waals forces	Electron gas
Properties	Hard; high melting points; may be soluble in polar liquids such as water; electrical insulators (but conductors in solution)	Very hard; high melting points; insoluble in nearly all liquids; semiconductors (except diamond, which is an insulator)	Soft; low melting and boiling points; soluble in covalent liquids; electrical insulators	Ductile; metallic luster; high electrical and thermal conductivity
Example	Sodium chloride, NaCl $E_{\text{cohesive}} = 3.28 \text{ eV/atom}$	Diamond, C $E_{\text{cohesive}} = 7.4 \text{ eV/atom}$	Methane, CH <sub>4</sub> $E_{\text{cohesive}} = 0.1 \text{ eV/molecule}$	Sodium, Na $E_{\text{cohesive}} = 1.1 \text{ eV/atom}$

Metallic bonding occurs when the reduction in electron potential energy outbalances the increase in electron KE that accompanies it. The more valence electrons per atom, the higher the average KE of the free electrons, but without a commensurate drop in their potential energy. For this reason nearly all the metallic elements are found in the first three groups of the periodic table.

### Ohm's Law

When the potential difference across the ends of a metal conductor is  $V$ , the resulting current  $I$  is, within wide limits, directly proportional to  $V$ . This empirical observation, called Ohm's law, is usually expressed as

$$\text{Ohm's law} \quad I = \frac{V}{R} \quad (10.9)$$

Here  $R$ , the resistance of the conductor, depends on its dimensions, composition, and temperature, but is independent of  $V$ . Ohm's law follows from the free-electron model of a metal.

We begin by assuming that the free electrons in a metal, like the molecules in a gas, move in random directions and undergo frequent collisions. The collisions here, however, are not billiard-ball collisions with other electrons but represent the scattering of electron waves by irregularities in the crystal structure, both defects such as impurity atoms and also atoms temporarily out of place as they vibrate. As we will see later, the atoms of a perfect crystal lattice do not scatter free electron waves except under certain specific circumstances.

If  $\lambda$  is the mean free path between the collisions of a free electron, the average time  $\tau$  between collisions is

$$\text{Collision time} \quad \tau = \frac{\lambda}{v_F} \quad (10.10)$$

The quantity  $v_F$  is the electron velocity that corresponds to the Fermi energy  $\epsilon_F$ , since only electrons at or near the top of their energy distribution can be accelerated (see Sec. 9.10). This average time is virtually independent of an applied electric field  $E$  because  $v_F$  is extremely high compared with the velocity change such a field produces. In copper, for instance,  $\epsilon_F = 7.04 \text{ eV}$  and so

$$v_F = \frac{2\epsilon_F}{m} = \sqrt{\frac{(2)(7.04 \text{ eV})(1.60 \times 10^{-19} \text{ J/eV})}{9.11 \times 10^{-31} \text{ kg}}} = 1.57 \times 10^6 \text{ m/s}$$

The superimposed drift velocity  $v_d$  due to an applied electric field, however, is usually less than 1 mm/s.

### Example 10.2

Find the drift velocity  $v_d$  of the free electrons in a copper wire whose cross-sectional area is  $A = 1.0 \text{ mm}^2$  when the wire carries a current of 1.0 A. Assume that each copper atom contributes one electron to the electron gas.

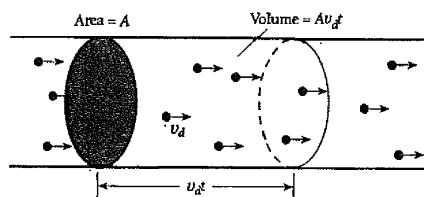


Figure 10.16 The number of free electrons in a wire that drift past a cross-section of the wire in the time  $t$  is  $nV = nAv_d t$ , where  $n$  is the number of free electrons/m<sup>3</sup> in the wire.

#### Solution

The wire contains  $n$  free electrons per unit volume. Each electron has the charge  $e$  and in the time  $t$  it travels the distance  $v_d t$  along the wire, as in Fig. 10.16. The number of free electrons in the volume  $Av_d t$  is  $nAv_d t$ , and all of them pass through any cross section of the wire in the time  $t$ . Thus the charge that passes through this cross section in  $t$  is  $Q = nAev_d t$ , and the corresponding current is

$$I = \frac{Q}{t} = nAev_d$$

The drift velocity of the electrons is therefore

$$v_d = \frac{I}{nAe}$$

From Example 9.8 we know that, in copper,  $n = N/V = 8.5 \times 10^{28}$  electrons/m<sup>3</sup>, and here  $I = 1.0$  A and  $A = 1.0$  mm<sup>2</sup> =  $1.0 \times 10^{-6}$  m<sup>2</sup>. Hence

$$v_d = \frac{1.0\text{ A}}{(8.5 \times 10^{28} \text{ m}^{-3})(1.0 \times 10^{-6} \text{ m}^2)(1.6 \times 10^{-19} \text{ C})} = 7.4 \times 10^{-4} \text{ m/s}$$

But if the free electrons have so small a drift velocity, why does an electric appliance go on as soon as its switch is closed and not minutes or hours later? The answer is that applying a potential difference across a circuit very rapidly creates an electric field in the circuit, and as a result all the free electrons begin their drift almost simultaneously.

A potential difference  $V$  across the ends of a conductor of length  $L$  produces an electric field of magnitude  $E = V/L$  in the conductor. This field exerts a force of  $eE$  on a free electron in the conductor, whose acceleration is

$$a = \frac{F}{m} = \frac{eE}{m} \quad (10.11)$$

When the electron undergoes a collision, it rebounds in an arbitrary direction and, on the average, no longer has a component of velocity parallel to  $E$ . Imposing the field  $E$  on the free electron gas in a metal superimposes a general drift on the faster but random motions of the electron (Fig. 10.17). We can therefore ignore the electron's motion at the Fermi velocity  $v_F$  in calculating the drift velocity  $v_d$ .

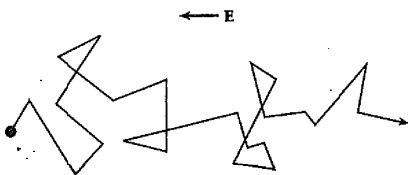


Figure 10.17 An electric field produces a general drift superimposed on the random motion of a free electron. The electron's path between collisions is actually slightly curved because of the acceleration due to the field.

After each collision, the electron is accelerated for some time interval  $\Delta t$  before the next collision, and at the end of the interval has traveled  $\frac{1}{2}a \Delta t^2$ . When the electron has made many collisions, its average displacement will be  $\bar{X} = \frac{1}{2}a \bar{\Delta t}^2$ , where  $\bar{\Delta t}^2$  is the average of the squared time intervals. Because of the way  $\Delta t$  varies,  $\bar{\Delta t}^2 = 2\tau^2$ . Hence  $\bar{X} = a\tau^2$  and the drift velocity is  $\bar{X}/\tau = a\tau$ , so that

$$\text{Drift velocity } v_d = a\tau = \left(\frac{eE}{m}\right)\left(\frac{\lambda}{v_F}\right) = \frac{eE\lambda}{mv_F} \quad (10.12)$$

In Example 10.2 we found that the current  $I$  in a conductor of cross-sectional area  $A$  in which the free electron density is  $n$  is given by

$$I = nAev_d \quad (10.13)$$

Using the value of  $v_d$  from Eq. (10.12) gives

$$I = \frac{nAe^2 E \lambda}{mv_F}$$

Since the electric field in the conductor is  $E = V/L$ ,

$$I = \left(\frac{n e^2 \lambda}{m v_F}\right) \left(\frac{A}{L}\right) V \quad (10.14)$$

This formula becomes Ohm's law if we set

$$\text{Resistance of metal conductor} \quad R = \left(\frac{m v_F}{n e^2 \lambda}\right) \frac{L}{A} \quad (10.15)$$

The quantity in parentheses is known as the resistivity  $\rho$  of the metal and is a constant for a given sample at a given temperature:

$$\text{Resistivity} \quad \rho = \frac{m v_F}{n e^2 \lambda} \quad (10.16)$$

**Example 10.3**

The resistivity of copper at 20°C is  $\rho = 1.72 \times 10^{-8} \Omega \cdot \text{m}$ . Estimate the mean free path  $\lambda$  between collisions of the free electrons in copper at 20°C.

**Solution**

In Example 9.8 we found that the free electron density in copper is  $n = 8.48 \times 10^{28} \text{ m}^{-3}$ , and earlier in this section we saw that the Fermi velocity there is  $v_F = 1.57 \times 10^6 \text{ m/s}$ . Solving Eq. (10.16) for  $\lambda$  gives

$$\lambda = \frac{mv_F}{ne^2\rho} = \frac{(9.11 \times 10^{-31} \text{ kg})(1.57 \times 10^6 \text{ m/s})}{(8.48 \times 10^{28} \text{ m}^{-3})(1.60 \times 10^{-19} \text{ C})^2(1.72 \times 10^{-8} \Omega \cdot \text{m})} \\ = 3.83 \times 10^{-8} \text{ m} = 38.3 \text{ nm}$$

The ions in solid copper are 0.26 nm apart, so a free electron travels past nearly 150 of them, on the average, before being scattered.

The scattering of free electron waves in a metal that leads to its electric resistance is caused both by structural defects and by ions out of place as they vibrate. Imperfections of the former kind do not depend on temperature but on the purity of the metal and on its history. The resistivities of cold-worked metals (such as "hard drawn" wires) are lowered by annealing because the number of defects is thereby decreased. On the other hand, lattice vibrations increase in amplitude with increasing temperature, and their contribution to resistivity accordingly goes up with temperature. Thus the resistivity of a metal is the sum  $\rho = \rho_t + \rho_b$  where  $\rho_t$  depends on the concentration of defects and  $\rho_b$  depends on temperature.

Figure 10.18 shows how the resistivities of two sodium samples vary with temperature. The top curve corresponds to the sample with the higher concentration of defects, which accounts for its upward displacement. In very pure and almost defect-free samples,  $\rho_t$  is small, and at low temperatures,  $\rho_t$  is also small. When both these conditions hold in copper, for example, the mean free path may be 10<sup>5</sup> times the value found in Example 10.3.

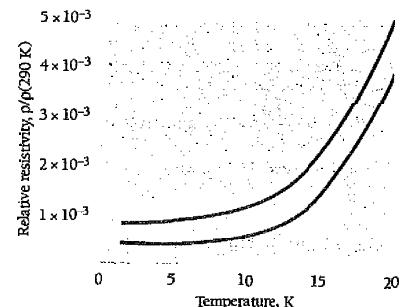


Figure 10.18 Resistivities of two sodium samples at low temperatures relative to their resistivities at 290 K. The upper curve corresponds to the sample with the higher concentration of impurities.

### Weidemann-Franz Law

The free-electron model of metallic conduction was proposed by Paul Drude in 1900, only three years after the discovery of the electron by J. J. Thomson, and was later elaborated by Hendrik Lorentz. Fermi-Dirac statistics were unknown then, and Drude and Lorentz assumed that the free electrons were in thermal equilibrium with a Maxwell-Boltzmann velocity distribution. This meant that the  $v_F$  in Eq. (10.16) was replaced by the rms electron velocity  $v_{rms}$ . In addition, Drude and Lorentz assumed that the free electrons collide with the metal ions, not with the much farther apart lattice defects. The net result was resistivity values on the order of 10 times greater than the measured ones.

The theory was nevertheless considered to be on the right track, both because it gave the correct form of Ohm's law and also because it accounted for the Weidemann-Franz law. This empirical law states that the ratio  $K/\sigma$  (where  $\sigma = 1/\rho$ ) between thermal and electric conductivities is the same for all metals and is a function only of temperature. If there is a temperature difference  $\Delta T$  between the sides of a slab of material  $\Delta x$  thick whose cross-sectional area is  $A$ , the rate  $\Delta Q/\Delta t$  at which heat passes through the slab is given by

$$\frac{\Delta Q}{\Delta t} = -KA \frac{\Delta T}{\Delta x}$$

where  $K$  is the thermal conductivity. According to the kinetic theory of a classical gas applied to the electron gas in the Drude-Lorentz model,

$$K = \frac{k n v_{rms} \lambda}{2}$$

From Eq. (10.16) with  $v_F$  replaced by  $v_{rms}$ ,

$$\sigma = \frac{1}{\rho} = \frac{n e^2 \lambda}{m v_{rms}}$$

Hence the ratio between the thermal and electric resistivities of a metal is

$$\frac{K}{\sigma} = \left( \frac{k n v_{rms} \lambda}{2} \right) \left( \frac{m v_{rms}}{n e^2 \lambda} \right) = \frac{k m v_{rms}^2}{2 e^2}$$

According to Eq. (9.15),  $v_{rms}^2 = 3kT/m$ , which gives

$$\frac{K}{\sigma T} = \frac{3k^2}{2e^2} = 1.11 \times 10^{-8} \text{ W} \cdot \Omega / \text{K}^2$$

This ratio does not contain the electron density  $n$  or the mean free path  $\lambda$ , so  $K/\sigma T$  ought to have the same constant value for all metals, which is the Weidemann-Franz law. To be sure, the above value of  $K/\sigma T$  is incorrect because it is based on a Maxwell-Boltzmann distribution of electron velocities. When Fermi-Dirac statistics are used, the result is

$$\frac{K}{\sigma T} = \frac{\pi^2 k^2}{3e^2} = 2.45 \times 10^{-8} \text{ W} \cdot \Omega / \text{K}^2$$

which agrees quite well with experimental findings.



## 10.6 BAND THEORY OF SOLIDS

*The energy band structure of a solid determines whether it is a conductor, an insulator, or a semiconductor*

No property of solids varies as widely as their ability to conduct electric current. Copper, a good conductor, has a resistivity of  $\rho = 1.7 \times 10^{-8} \Omega \cdot \text{m}$  at room temperature,



Felix Bloch (1905–1983) was born in Zurich, Switzerland, and did his undergraduate work in engineering there. He went to Leipzig in Germany for his Ph.D. in physics, remaining there until the rise of Hitler. In 1934 Bloch joined the faculty of Stanford University where he stayed until his retirement except for the war years, which he spent at Los Alamos helping develop the atomic bomb, and for 1954 to 1955, when he was the first director of CERN, the European center for nuclear and elementary-particle research in Geneva.

In 1928 in his doctoral thesis Bloch showed how allowed and forbidden bands arise by solving Schrödinger's equation for an electron moving in the periodic potential of a crystal. This important step in the development of the theory of solids supplemented earlier work by Walter Heitler and Fritz London, who showed how energy levels broaden into bands when atoms are brought together to form a solid. Later Bloch studied the magnetic behavior of atomic nuclei in solids and liquids, which led to the extremely sensitive nuclear magnetic resonance method of analysis. Bloch received the Nobel Prize in physics in 1952 together with Edward Purcell of Harvard, who had also done important work in nuclear magnetism.

whereas for quartz, a good insulator,  $\rho = 7.5 \times 10^{17} \Omega \cdot \text{m}$ , more than 25 powers of ten greater. The existence of electron energy bands in solids makes it possible to understand this remarkable span.

There are two ways to consider how energy bands arise. The simplest is to look at what happens to the energy levels of isolated atoms as they are brought closer and closer together to form a solid. We will begin in this way and then examine the significance of energy bands. Later we will consider the origin of energy bands in terms of the restrictions the periodicity of a crystal lattice imposes on the motion of electrons.

The atoms in every solid, not just in metals, are so near one another that their valence electron wave functions overlap. In Sec. 8.3 we saw the result when two H atoms are brought together. The original 1s wave functions can combine to form symmetric or antisymmetric joint wave functions, as in Figs. 8.5 and 8.6, whose energies are different. The splitting of the 1s energy level in an isolated H atom into two levels, marked  $E_A^{\text{total}}$  and  $E_S^{\text{total}}$ , is shown as a function of internuclear distance in Fig. 8.7.

The greater the number of interacting atoms, the greater the number of levels produced by the mixing of their respective valence wave functions (Fig. 10.19). In a solid, because the splitting is into as many levels as there are atoms present (nearly  $10^{23}$  in a cubic centimeter of copper, for instance), the levels are so close together that they form an energy band that consists of a virtually continuous spread of permitted energies. The energy bands of a solid, the gaps between them, and the extent to which they are filled by electrons not only govern the electrical behavior of the solid but also have important bearing on others of its properties.

### Conductors

Figure 10.20 shows the energy levels and bands in sodium. The 3s level is the first occupied level to broaden into a band. The lower 2p level does not begin to spread out until a much smaller internuclear distance because the 2p wave functions are closer to the nucleus than are the 3s wave functions. The average energy in the 3s band drops at first, which signifies attractive forces between the atoms. The actual internuclear distance in solid sodium corresponds to the minimum average 3s electron energy.

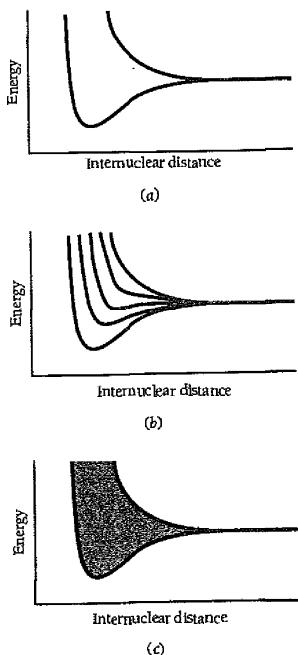


Figure 10.19 The 3s level is the highest occupied level in a ground-state sodium atom. (a) When two sodium atoms come close together, their 3s levels, initially equal, become two separate levels because of the overlap of the corresponding electron wave functions. (b) The number of new levels equals the number of interacting atoms, here 5. (c) When the number of interacting atoms is very large, as in solid sodium, the result is an energy band of very closely spaced levels.

An electron in a solid can only have energies that fall within its energy bands. The various outer energy bands in a solid may overlap, as in Fig. 10.21a, in which case its valence electrons have available a continuous distribution of permitted energies. In other solids the bands may not overlap, as in Fig. 10.21b, and the intervals between them represent energies their electrons cannot have. Such intervals are called forbidden bands or band gaps.

Figure 9.11 shows the distribution of electron energies in a band at various temperatures. At 0 K all levels in the band are filled by electrons up to the Fermi energy  $\epsilon_F$ , and those above  $\epsilon_F$  are empty. At temperatures above 0 K, electrons with energies below  $\epsilon_F$  can move into higher states, in which case  $\epsilon_F$  represents a level with a 50 percent likelihood of being occupied.

A sodium atom has a single 3s valence electron. Each s ( $l = 0$ ) atomic level can hold  $2(2l + 1) = 2$  electrons, so each s band formed by  $N$  atoms can hold  $2N$  electrons.

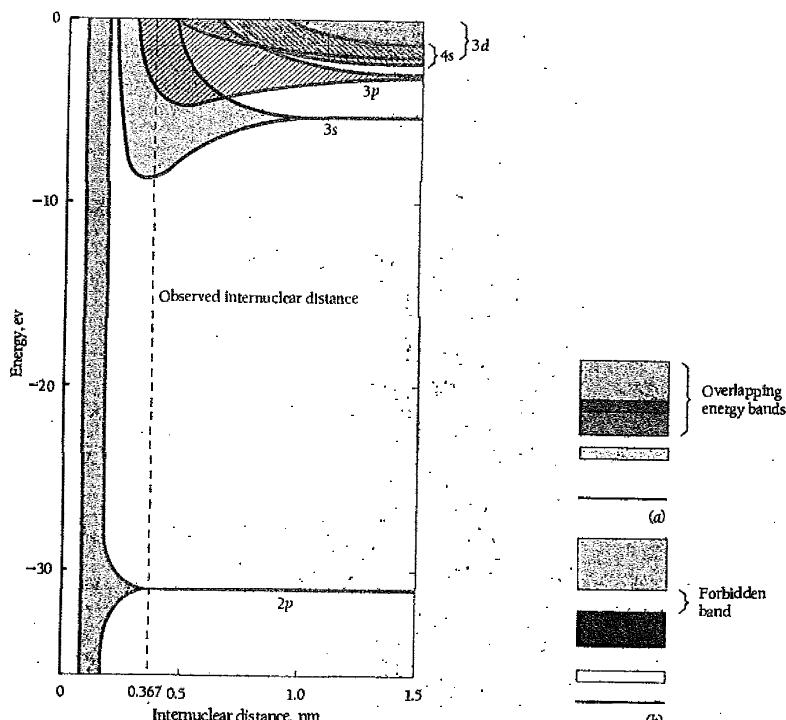


Figure 10.20 The energy levels of sodium atoms become bands as their internuclear distance decreases. The observed internuclear distance in solid sodium is 0.367 nm.

Thus the 3s band in solid sodium is only half filled by electrons (Fig. 10.22) and the Fermi energy  $\epsilon_F$  lies in the middle of the band.

When a potential difference is applied across a piece of solid sodium, 3s electrons can pick up additional energy while remaining in their original band. The additional energy is in the form of KE, and the drift of the electrons constitutes an electric current. Sodium is therefore a good conductor, as are other solids with partly filled energy bands.

Magnesium atoms have filled 3s shells. If the 3s level simply spreads into a 3s band in solid magnesium, as in Fig. 10.21b, there would be a forbidden band above it and the 3s electrons could not easily pick up enough energy to jump the forbidden band to the empty band above it. Nevertheless magnesium is a metal. What actually happens is that the 3p and 3s bands overlap as magnesium atoms become close together to give the structure shown in Fig. 10.21a. A  $p$  ( $l = 1$ ) atomic level can hold  $2(2l + 1) = 2(2 + 1) = 6$  electrons, so a  $p$  band formed by  $N$  atoms can hold  $6N$  electrons. Together

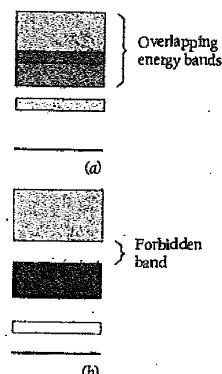


Figure 10.21 (a) The energy bands in some solids may overlap to give a continuous band. (b) A forbidden band separate from overlapping energy bands in other solids.

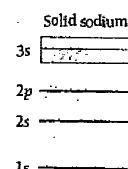


Figure 10.22 The 3s energy band in solid sodium is half filled with electrons. The Fermi energy  $\epsilon_F$  is in the middle of the band.

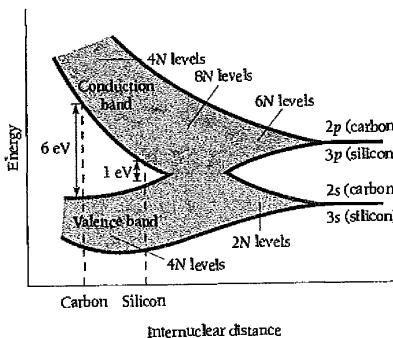


Figure 10.23 Origin of the energy bands of carbon and silicon. The 2s and 2p levels of carbon atoms and the 3s and 3p levels of silicon atoms spread into bands that first overlap with decreasing atomic separation and then split into two diverging bands. The lower band is occupied by valence electrons and the upper conduction band is empty. The energy gap between the bands depends on the internuclear separation and is greater for carbon than for silicon.

with the 2N electrons the 3s band can hold, the  $3s + 3p$  band in magnesium can hold  $8N$  electrons in all. With only  $2N$  electrons in the band, it is only one-quarter filled and so magnesium is a conductor.

### Insulators

In a carbon atom the  $2p$  shell contains only two electrons. Because a  $p$  shell can hold six electrons, we might think that carbon is a conductor, just as sodium is. What actually happens is that, although the  $2s$  and  $2p$  bands that form when carbon atoms come together overlap at first (as the  $3s$  and  $3p$  bands in sodium do), at smaller separations the combined band splits into two bands (Fig. 10.23), each able to contain  $4N$  electrons. Because a carbon atom has two  $2s$  and two  $2p$  electrons, in diamond there are  $4N$  valence electrons that completely fill the lower (or valence) band, as in Fig. 10.24. The empty conduction band above the valence band is separated from it by a forbidden band  $6\text{ eV}$  wide. Here the Fermi energy  $\epsilon_F$  is at the top of the valence band. At least  $6\text{ eV}$  of additional energy must be provided to an electron in diamond

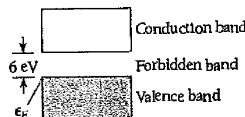


Figure 10.24 Energy bands in diamond. The Fermi energy is at the top of the filled lower band. Because an electron in the valence band needs at least  $6\text{ eV}$  to reach the empty conduction band, diamond is an insulator.



Figure 10.25 The valence and conduction bands in a semiconductor are separated by a smaller gap than in the case of an insulator. Here a small number of electrons near the top of the valence band can acquire enough thermal energy to jump the gap and enter the conduction band. The Fermi energy is therefore in the middle of the gap.

if it is to climb to the conduction band where it can move about freely. With  $kT = 0.025$  eV at room temperature, valence electrons in diamond do not have enough thermal energy to jump the 6 eV gap.

Nor can an energy increment of 6 eV be given to a valence electron in diamond by an electric field, because such an electron undergoes frequent collisions with crystal imperfections during which it loses most of the energy it gains from the field. An electric field of over  $10^8$  V/m is needed for an electron to gain 6 eV in a typical mean free path of  $5 \times 10^{-8}$  m. This is billions of times stronger than the field needed for a current to flow in a metal. Diamond is therefore a very poor conductor and is classed as an insulator.

### Semiconductors

Silicon has a crystal structure like that of diamond and, as in diamond, a gap separates the top of its filled valence band from an empty conduction band above it (see Fig. 10.23). The forbidden band in silicon, however, is only about 1 eV wide. At low temperatures silicon is little better than diamond as a conductor, but at room temperature a small number of its valence electrons have enough thermal energy to jump the forbidden band and enter the conduction band (Fig. 10.25). These electrons, though few, are still enough to allow a small amount of current to flow when an electric field is applied. Thus silicon has a resistivity intermediate between those of conductors and those of insulators, and it and other solids with similar band structures are classed as semiconductors.

### Impurity Semiconductors

Small amounts of impurity can drastically change the conductivity of a semiconductor. Suppose we incorporate a few arsenic atoms in a silicon crystal. Arsenic atoms have five electrons in their outer shells, silicon atoms have four. (These shells have the configurations  $4s^24p^3$  and  $3s^23p^2$  respectively.) When an arsenic atom replaces a silicon atom in a silicon crystal, four of its electrons participate in covalent bonds with its nearest neighbors. The fifth electron needs very little energy—only about 0.05 eV in silicon, about 0.01 eV in germanium—to be detached and move about freely in the crystal.

As shown in Fig. 10.26, arsenic as an impurity in silicon provides energy levels just below the conduction band. Such levels are called donor levels, and the substance is called an *n*-type semiconductor because electric current in it is carried by negative charges (Fig. 10.27). The presence of donor levels below the conduction band raises the Fermi energy above the middle of the forbidden band between the valence and conduction bands.

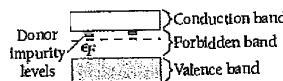


Figure 10.26 A trace of arsenic in a silicon crystal provides donor levels in the normally forbidden band, producing an *n*-type semiconductor.

### Optical Properties of Solids

The optical properties of solids are closely related to their energy-band structures. Photons of visible light have energies from about 1 to 3 eV. A free electron in a metal can readily absorb such an amount of energy without leaving its valence band, and metals are accordingly opaque. The characteristic luster of a metal is due to the reradiation of light absorbed by its free electrons. If the metal surface is smooth, the reradiated light appears as a reflection of the original incident light.

For a valence electron in an insulator to absorb a photon, on the other hand, the photon energy must be over 3 eV if the electron is to jump across the forbidden band to the conduction band. Insulators therefore cannot absorb photons of visible light and are transparent. Of course, most samples of insulating materials do not appear transparent, but this is due to the scattering of light by irregularities in their structures. Insulators are opaque to ultraviolet light, whose higher frequencies mean high enough photon energies to allow electrons to cross the forbidden band.

Because the forbidden bands in semiconductors are about the same in width as the photon energies of visible light, they are usually opaque to visible light. However, they are transparent to infrared light whose lower frequencies mean photon energies too low to be absorbed. For this reason infrared lenses can be made from the semiconductor germanium, whose appearance in visible light is that of an opaque solid.

If we instead incorporate gallium atoms in a silicon crystal, a different effect occurs. Gallium atoms have only three electrons in their outer shells, whose configuration is  $4s^24p$ , and their presence leaves vacancies called holes in the electron structure of the crystal. An electron needs relatively little energy to enter a hole, but as it does so, it leaves a new hole in its former location. When an electric field is applied across a silicon crystal containing a trace of gallium, electrons move toward the anode by successively filling holes (Fig. 10.28). The flow of current here is conveniently described with reference to the holes, whose behavior is like that of positive charges since they move toward the negative electrode. A substance of this kind is called a *p*-type semiconductor.

In the energy-band diagram of Fig. 10.29 we see that gallium as an impurity in silicon provides energy levels, called acceptor levels, just above the valence band. Any electrons that occupy these levels leave behind them vacancies in the valence band that permit electric current to flow. The Fermi energy in a *p*-type semiconductor lies below the middle of the forbidden band.

Adding an impurity to a semiconductor is called doping. Phosphorus, antimony, and bismuth as well as arsenic have atoms with five valence electrons and so can be used as donor impurities in doping silicon and germanium to yield an *n*-type semiconductor. Similarly, indium and thallium as well as gallium have atoms with three valence electrons and so can be used as acceptor impurities. A minute amount of impurity can produce a dramatic change in the conductivity of a semiconductor. As an example, 1 part

Figure 10.27 Current in an *n*-type semiconductor is carried by surplus electrons that do not fit into the electron structure of a pure crystal.

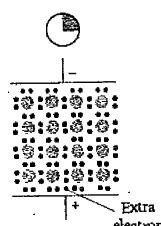
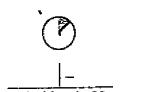
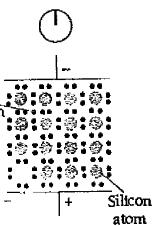




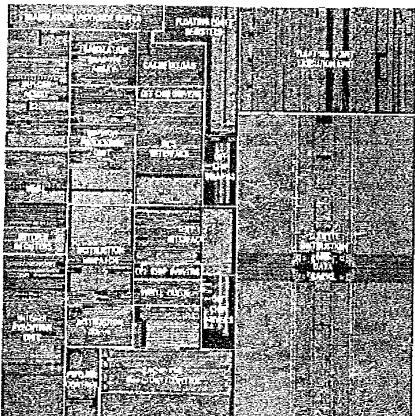
Figure 10.29 A trace of gallium in a silicon crystal provides acceptor levels in the normally forbidden band, producing a p-type semiconductor.

of a donor impurity per  $10^9$  parts of germanium increases its conductivity by a factor of nearly  $10^3$ . Silicon and germanium are not the only semiconducting materials with practical applications; another important class of semiconductors consists of compounds of trivalent and pentavalent elements, such as GaAs, GaP, InSb, and InP.

## 10.7 SEMICONDUCTOR DEVICES

*The properties of the p-n junction are responsible for the microelectronics industry*

The operation of most semiconductor devices is based upon the nature of junctions between p- and n-type materials. Such junctions can be made in several ways. A method especially adapted to the production of integrated circuits involves diffusing impurities in vapor form into a semiconductor wafer in regions defined by masks. A series of diffusion steps using donor and acceptor impurities is part of the procedure for manufacturing circuits that can contain millions of resistors, capacitors, diodes, and transistors on a chip a few millimeters across. The limiting factor in this method is the wavelength of the light that is shined through masks to expose and thereby harden the



The IBM PowerPC 601 microprocessor chip is 10.95 mm square and contains 2.8 million transistors. The functions of the various parts of the chip are indicated.

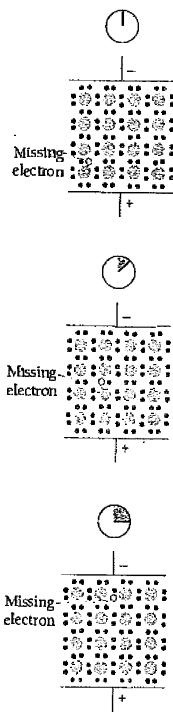


Figure 10.28 Current in a p-type semiconductor is carried by the motion of "holes," which are sites of missing electrons. Holes move toward the negative electrode as a succession of electrons move into them.

photoresist compound on the wafer surface. (The unexposed photoresist is then washed away to leave areas open to the next diffusion step.) The shortest wavelength that can be used with conventional optical systems is 193 nm (which is in the ultraviolet) because no suitable material transparent to shorter wavelengths is known that can be made into lenses. Features as small as 130 nm might possibly be created with 193-nm light, but the demands of the electronics industry for ever-more components per chip will soon have to be met with some other technology. X-rays, electron and ion beams are being studied for this purpose, with an immediate goal of chips with 200 million circuit elements each 100 nm across.

### Junction Diode

A characteristic property of a *p-n* junction is that electric current can pass through it much more readily in one direction than in the other. In the diode shown in Fig. 10.30, the left-hand end is a *p*-type region in which conduction involves the motion of holes, and the right-hand end is an *n*-type region in which conduction occurs by means of the motion of electrons. Three situations can occur:

1 **No bias** This is illustrated in Fig. 10.30a. Electron-hole pairs are created spontaneously by thermal excitation in the valence band of the *p*-region. Some of the electrons have enough energy to jump the gap to the conduction band and then migrate to the *n* region. There they lose energy in collisions. At the same time, some electrons in the *n* region are sufficiently energetic to climb the energy hill and enter the *p* region, where they recombine with holes there. At thermal equilibrium the two processes occur at the same low rate, so there is no net current. The Fermi energy is the same in both *p* and *n* regions; if it were not, electrons would flow to the region with vacant states of lower energy until  $\epsilon_F$  is the same.

2 **Reverse bias** As in Fig. 10.30b, an external voltage  $V$  is applied across the diode with the *p* end negative and the *n* end positive. The energy difference across the junction is greater by  $Ve$  than in part a, which impedes the recombination current  $i_r$ ; the holes in the *p* region migrate to the left and are filled at the negative terminal, while the electrons in the *n* region migrate to the right and leave the diode at the positive terminal. New electron-hole pairs are still being created as before by thermal excitation, but because they are relatively few in number the resulting net current  $i_g - i_r$  is very small even when the applied voltage  $V$  is high. (We note that the conventional current  $I$ , which flows from + to -, is opposite in direction to the electron current  $i$ .)

3 **Forward bias** As in Fig. 10.30c, the external voltage is applied with the *p* end of the diode positive and the *n* end negative. The energy difference across the junction is now less by  $Ve$  than in part a, which increases the recombination current  $i_r$ , since the electrons have a smaller energy hill to climb. Under these circumstances new holes are created continuously by the removal of electrons at the positive terminal while new electrons are added at the negative terminal. The holes migrate to the right and the electrons to the left under the influence of the applied potential. The holes and electrons meet in the vicinity of the *p-n* junction and recombine there.

Thus current can flow readily in one direction through a *p-n* junction but hardly at all in the other direction, which makes such a junction an ideal rectifier in an electric circuit. The greater the applied voltage, the greater the current in the forward direction. Figure 10.31 shows how  $I$  varies with  $V$  for a *p-n* junction rectifier.

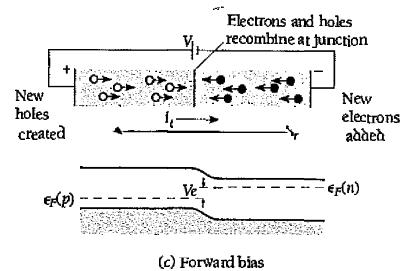
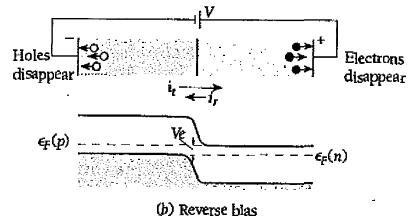
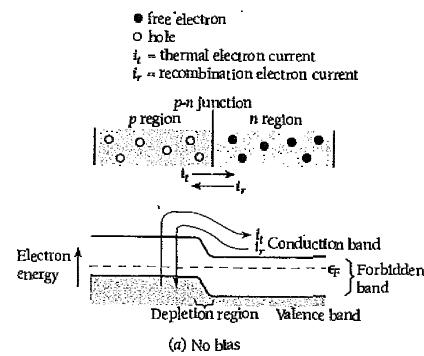


Figure 10.30 Operation of a semiconductor diode.

(a) When there is no applied voltage, the thermal electron current to the right equals the recombination electron current to the left and there is no net current. Both these currents are small.

(b) When an external voltage is applied so that the p end of the diode is negative, the recombination electron current is less than the thermal electron current. The result is a very small net electron current to the right.

(c) When an external voltage is applied so that the p end of the diode is positive, the recombination current can be much larger than the thermal electron current to give a large net electron current to the left. The conventional current is in the opposite direction to the electron current.

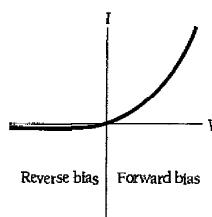


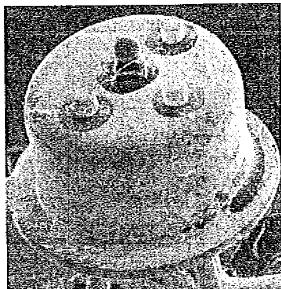
Figure 10.31 Voltage-current characteristic of a p-n semiconductor diode.

### Photodiodes

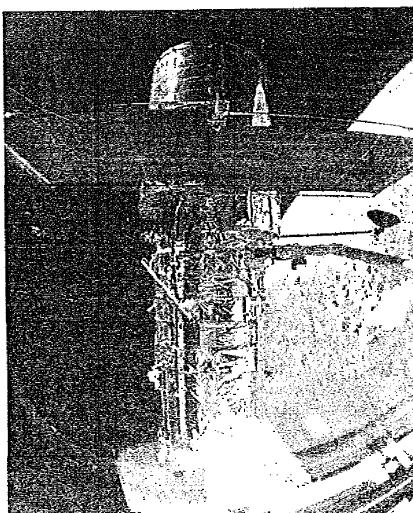
Energy is needed to create an electron-hole pair, and this energy is released when an electron and a hole recombine. In silicon and germanium the recombination energy is absorbed by the crystal as heat, but in certain other semiconductors, for instance gallium arsenide, a photon is emitted when recombination occurs. This is the basis of the light-emitting diode (LED). Forward bias is used in an LED, so the electrons and holes both move toward the *p-n* junction, as in Fig. 10.30c, where they recombine to create photons.

A fairly small current is used in an LED and the photons are produced by spontaneous emission. When the current is high, spontaneous emission may not keep up with the rate of arrival of electrons and holes in the depletion region, and the result is a substantial population inversion there. This is the condition for laser action to occur, with spontaneously emitted photons causing avalanches of additional photons by stimulated emission. In a semiconductor laser opposite ends of the *p-n* junction are made parallel and partly reflecting. The coherent light produced by the stimulated emission is intensified as it moves back and forth in the thin depletion region, and emerges through the ends (Fig. 10.32).

The process that occurs in an LED is reversed in a silicon solar cell. Here photons arriving at or near the depletion region of a *p-n* junction after passing through a thin ( $<1\text{ }\mu\text{m}$ ) outer layer of silicon produce electron-hole pairs if sufficiently energetic. The electrons are raised to the conduction band, leaving holes in the valence band. The potential difference across the depletion region provides an electric field that pulls the electrons to the *n* region and the holes to the *p* region. The newly freed electrons can then flow from the *n* region through an external



This light-emitting diode has a spherical glass lens mounted on it. The diode is made of gallium arsenide doped with phosphorus and produces monochromatic red light of wavelength 620 nm for use with a fiber-optic telephone transmission line.



The Hubble Space Telescope being launched from the Space Shuttle Discovery. One of the two arrays of solar cells that power the telescope has been deployed.

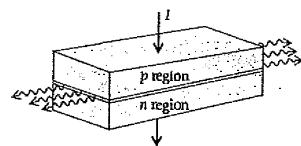


Figure 10.32 A semiconductor laser. Each dimension is less than a millimeter and its light output, as in all lasers, is coherent. The junction between the *p* and *n* regions from which the light emerges is only a few micrometers thick.

circuit to the *p* region where they recombine with the newly created holes. In this way the energy of incident photons can be converted to electric energy. Diodes of this kind are widely used to detect photons in such devices as light meters in cameras as well as to produce electric energy from solar radiation.

The only charge carriers shown in Fig. 10.30 were the electrons. Actually of course, what was said also applies to the holes, which act as positive charges and behave in exactly the opposite way to add their current to the conventional current.

When a *p* material joins an *n* material, a depletion region occurs between them instead of a sharp interface, as shown in the lower part of Fig. 10.30a. In this region electrons from the donor levels of the *n* material fill the holes of the acceptor levels of the *p* material, so that few charge carriers of either kind are present there. The width of the depletion region depends on exactly how the diode is produced, and is typically about  $10^{-6}$  m.

#### Tunnel Diode

The *p* and *n* parts of a diode can be heavily doped to give the energy band structure of Fig. 10.33a. The depletion region is very narrow,  $\sim 10^{-8}$  m, and the bottom of the *n* conduction band overlaps the top of the *p* valence band. The large concentration of impurities causes the donor levels to merge into the bottom of the *n* conduction band, which moves the Fermi energy there upward into the band. Similarly the acceptor levels merge into the top of the *p* valence band, which lowers the Fermi energy below the top of the band.

Because the depletion region is so narrow, only a few electron wavelengths across, electrons can "tunnel" through the forbidden band there by the mechanism described in Sec. 5.9. For this reason such a diode is called a tunnel diode. When no external voltage is applied to the diode, electrons tunnel in both directions across the gap in equal numbers and the Fermi energy is constant across the diode.

Figure 10.33b shows what happens when a small forward voltage is applied to the diode. Now the filled lower part of the *n* conduction band is opposite the empty upper part of the *p* valence band, and the tunneling is from *n* to *p* only. This gives an electron current to the left, which corresponds to a conventional current to the right.

When the external voltage is increased further, the two bands no longer overlap, as in Fig. 10.33c. The tunnel current therefore ceases. From now on the diode behaves exactly like the ordinary junction diode of Fig. 10.30. Figure 10.34 shows the voltage-current characteristic curve of a tunnel diode.

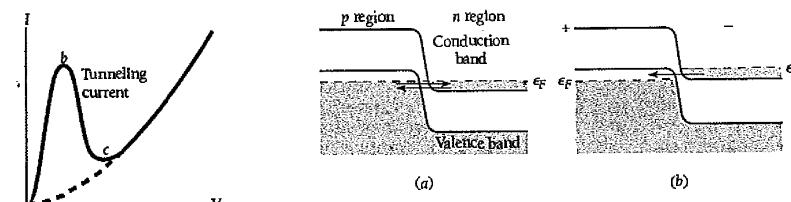


Figure 10.34 Voltage-current characteristic of a tunnel diode. The points *a*, *b*, and *c* correspond to parts *a*, *b*, and *c* of Fig. 10.33. The dashed line indicates the behavior of an ordinary junction diode, as in Fig. 10.30.

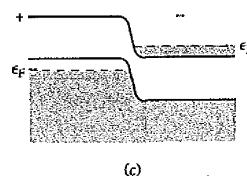


Figure 10.33 Operation of a tunnel diode. (a) No bias. Electrons tunnel both ways between the *p* and *n* regions. (b) Small forward bias. Electrons tunnel from the *n* to the *p* region only. (c) Larger forward bias. Now the valence band of the *p* region does not overlap the conduction band of the *n* region and so no tunneling can occur. At higher voltages the diode behaves like the ordinary diode of Fig. 10.30.

The importance of the tunnel diode lies in the rapidity with which a voltage change between *a* and *b* or between *b* and *c* in Fig. 10.34 can alter the current. In ordinary diodes and transistors, the response time depends on the diffusion speed of the charge carriers, which is low. Hence such devices operate slowly. Tunnel diodes, on the other hand, respond quickly to appropriate voltage changes and can be used in high-frequency oscillators and as fast switches in computers.

#### Zener Diode

Although the reverse current in many semiconductor diodes remains virtually constant even at high voltages, as in Fig. 10.31, in certain diodes the reverse current increases abruptly when a particular voltage is reached, as in Fig. 10.35. Such diodes are called Zener diodes and are widely used in voltage-regulation circuits.

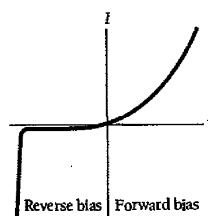


Figure 10.35 voltage-current characteristic of a Zener diode.

Two mechanisms contribute to the sharp rise in current. One, called avalanche multiplication, occurs when an electron near the junction is sufficiently accelerated by the electric field to ionize atoms it collides with, thereby creating fresh electron-hole pairs. The new electrons in their turn continue the process to produce a flood of charge carriers in the diode.

The other mechanism, called Zener breakdown, involves the tunneling of valence-band electrons on the *p* side of the junction to the conduction band on the *n* side even though these electrons do not have enough energy to first enter the conduction band on the *p* side. (Such tunneling is in the opposite direction to that occurring in a tunnel diode.) Zener breakdown can occur in heavily doped diodes at voltages of 6 V or less. In lightly doped diodes the necessary voltage is higher, and avalanche multiplication is then the chief process involved.

#### Junction Transistor

A transistor is a semiconductor device that can amplify a weak signal into a strong one when appropriately connected. Figure 10.36 shows an *n-p-n* junction transistor, which consists of a thin *p*-type region called the base that is sandwiched between two *n*-type regions called the emitter and the collector. (A *p-n-p* transistor behaves in a similar manner, except that the current then is carried by holes rather than by electrons.) The energy-band structure of an *n-p-n* transistor is given in Fig. 10.37.

The transistor is given a forward bias across the emitter-base junction and a reverse bias across the base-collector junction. The emitter is more heavily doped than the base, so nearly all the current across the emitter-base junction consists of electrons moving from left to right. Because the base is very thin ( $1 \mu\text{m}$  or so) and the concentration of holes there is low, most of the electrons entering the base diffuse through it to the base-collector junction where the high positive potential attracts them into the collector. Changes in the input-circuit current are thus mirrored by changes in the output-circuit current, which is only a few percent smaller.

The ability of the transistor of Fig. 10.36 to produce amplification comes about because the reverse bias across the base-collector junction permits a much higher voltage in the output circuit than that in the input circuit. Since electric power = (current)(voltage), the power of the output signal can greatly exceed the power of the input signal.

#### Field-Effect Transistor

Although its advent revolutionized electronics, the low input impedance of the junction transistor is a handicap in certain applications. In addition, it is difficult

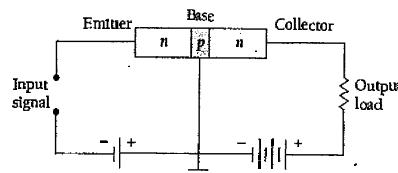
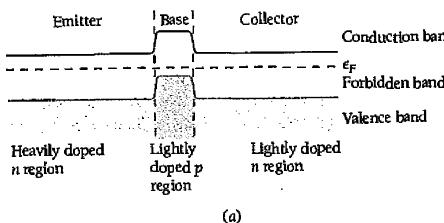
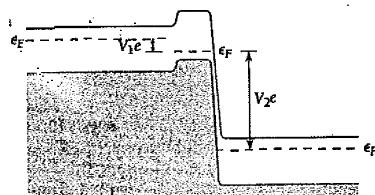


Figure 10.36 A simple junction-transistor amplifier.



(a)



(b)

Figure 10.37 (a) Isolated n-p-n transistor, (b) Transistor connected as in Fig. 10.36. The forward bias  $V_1$  between emitter and base is small; the reverse bias  $V_2$  between base and collector is large. Because the base is very thin, electrons can pass through it from emitter to collector without recombining with holes there. Once the electrons are in the collector, they undergo collisions in which they lose energy, and afterward cannot return to the base because the potential hill  $V_{2e}$  is too high.

to incorporate large numbers of them in an integrated circuit and they consume relatively large amounts of power. The field-effect transistor (FET) lacks these disadvantages and is widely used today although slower in operation than junction transistors.

As in Fig. 10.38, an n-channel FET consists of a block of n-type material with contacts at each end together with a strip of p-type material on one side that is called the gate. When connected as shown, electrons move from the source terminal to the drain terminal through the n-type channel. The p-n junction is given a reverse bias, and as a result both the n and p materials near the junction are depleted of charge carriers (see Fig. 10.30b). The higher the reverse potential on the gate, the larger the

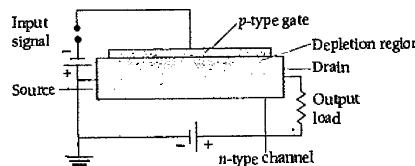


Figure 10.38 A field-effect transistor.

depleted region in the channel and the fewer the electrons available to carry the current. Thus the gate voltage controls the channel current. Very little current passes through the gate circuit owing to the reverse bias, and the result is an extremely high input impedance.

Even higher input impedances (up to  $10^{15} \Omega$ ) together with greater ease of manufacture are characteristic of the metal-oxide-semiconductor FET (MOSFET), a FET in which the semiconductor gate is replaced by a metal film separated from the channel by an insulating layer of silicon dioxide. The metal film is thus capacitively coupled to the channel, and its potential controls the drain current through the number of induced charges in the channel. A MOSFET occupies only a few percent of the area needed for a junction transistor.

## 10.8 ENERGY BANDS: ALTERNATIVE ANALYSIS

*How the periodicity of a crystal lattice leads to allowed and forbidden bands*

A very different approach can be taken to the origin of energy bands from that described in Sec. 10.6. There we saw that bringing together isolated atoms to form a solid has the effect of broadening their energy levels into bands of allowed electron energies. Alternatively we can start with the idea that an electron in a crystal moves in a region of periodically varying potential (Fig. 10.39) rather than one of constant potential. As a result diffraction effects occur that limit the electron to certain ranges of momenta that correspond to allowed energy bands. In this way of thinking, the interactions among the atoms influence the behavior of their valence electrons indirectly through the crystal lattice; these interactions bring about, rather than directly through the atomic interactions themselves. An intuitive approach will be used here to bring out more clearly the physics of the situation, instead of a formal treatment based on Schrödinger's equation.

The de Broglie wavelength of a free electron of momentum  $p$  is

$$\text{Free electron} \quad \lambda = \frac{\hbar}{p} \quad (10.17)$$

Unbound low-energy electrons can travel freely through a crystal since their wavelengths are long relative to the lattice spacing  $a$ . More energetic electrons, such as those with the Fermi energy in a metal, have wavelengths comparable with  $a$ , and such electrons are diffracted in precisely the same way as x-rays (Sec. 2.6) or electrons in a beam

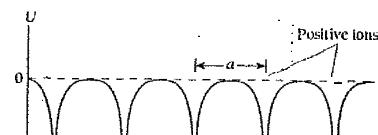


Figure 10.39 The potential energy of an electron in a periodic array of positive ions.

(Sec. 3.5) directed at the crystal from the outside. [When  $\lambda$  is near  $a$ ,  $2a$ ,  $3a$ , ... in length, Eq. (10.17) no longer holds, as discussed later.] An electron of wavelength  $\lambda$  undergoes Bragg reflection from one of the atomic planes in a crystal when it approaches the plane at an angle  $\theta$ , where from Eq. (2.13)

$$n\lambda = 2a \sin \theta \quad n = 1, 2, 3, \dots \quad (10.18)$$

It is customary to treat the situation of electron waves in a crystal by replacing  $\lambda$  by the wave number  $k$  introduced in Sec. 3.3, where

$$\text{Wave number} \quad k = \frac{2\pi}{\lambda} = \frac{p}{\hbar} \quad (10.19)$$

The wave number is equal to the number of radians per meter in the wave train it describes, and is proportional to the momentum  $p$  of the electron. Since the wave train moves in the same direction as the particle, we can describe the wave train by means of a vector  $\mathbf{k}$ . Bragg's formula in terms of  $k$  is

$$\text{Bragg reflection} \quad k = \frac{n\pi}{a \sin \theta} \quad n = 1, 2, 3, \dots \quad (10.20)$$

Figure 10.40 shows Bragg reflection in a two-dimensional square lattice. Evidently we can express the Bragg condition by saying that reflection from the vertical rows of ions occurs when the component of  $\mathbf{k}$  in the  $x$  direction,  $k_x$ , is equal to  $n\pi/a$ . Similarly, reflection from the horizontal rows occurs when  $k_y = n\pi/a$ .

Let us consider first electrons whose wave numbers are sufficiently small for them to avoid reflection. If  $k$  is less than  $\pi/a$ , the electron can move freely through the lattice in any direction. When  $k = \pi/a$ , they are prevented from moving in the  $x$  or  $y$  directions by reflection. The more  $k$  exceeds  $\pi/a$ , the more limited the possible directions of motion, until when  $k = \pi/a \sin 45^\circ = \sqrt{2}\pi/a$  the electrons are reflected, even when they move diagonally through the lattice.

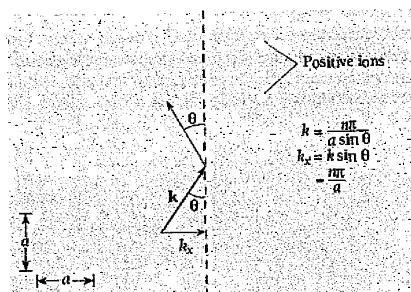


Figure 10.40 Bragg reflection from the vertical rows of ions occurs when  $k_x = n\pi/a$ .

### Brillouin Zones

The region in  $k$ -space (here an imaginary plane whose rectangular coordinates are  $k_x$  and  $k_y$ ) that low- $k$  electrons can occupy without being diffracted is called the first Brillouin zone, shown in Fig. 10.41. The second Brillouin zone is also shown; it contains electrons with  $k > \pi/a$  that do not fit into the first zone yet which have sufficiently small wave numbers to avoid diffraction by the diagonal sets of atomic planes in Fig. 10.40. The second zone contains electrons with  $k$  values from  $\pi/a$  to  $2\pi/a$  for electrons moving in the  $\pm x$  and  $\pm y$  directions, with the possible range of  $k$  values narrowing as the diagonal directions are approached. Further Brillouin zones can be constructed in the same manner. The extension of this analysis to actual three-dimensional structures leads to Brillouin zones such as those shown in Fig. 10.42.

The significance of the Brillouin zones becomes apparent when we look at the energies of the electrons in each zone.

The energy of a free electron is related to its momentum  $p$  by

$$\text{Energy and momentum} \quad E = \frac{p^2}{2m} \quad (10.21)$$

and hence to its wave number  $k$  by

$$\text{Energy and wave number} \quad E = \frac{\hbar^2 k^2}{2m} \quad (10.22)$$

In the case of an electron in a crystal for which  $k \ll \pi/a$ , there is practically no interaction with the lattice, and Eq. (10.22) is valid. Since the energy of such an electron

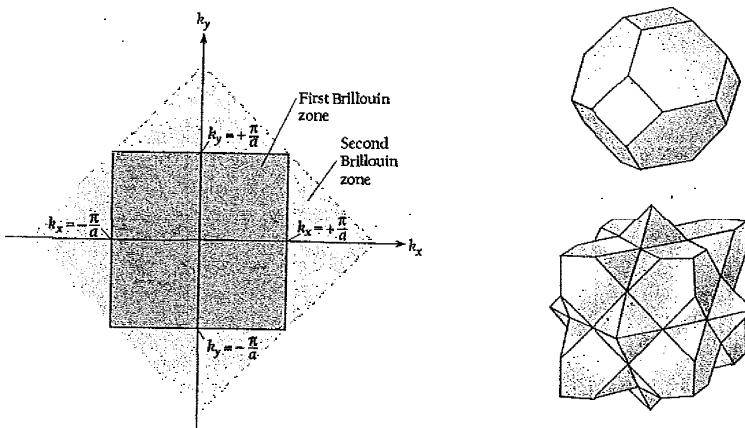
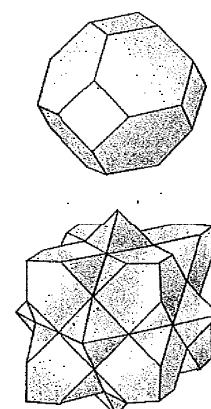


Figure 10.41 The first and second Brillouin zones of a two-dimensional square lattice.

Figure 10.42 First and second Brillouin zones in a face-centered crystal.



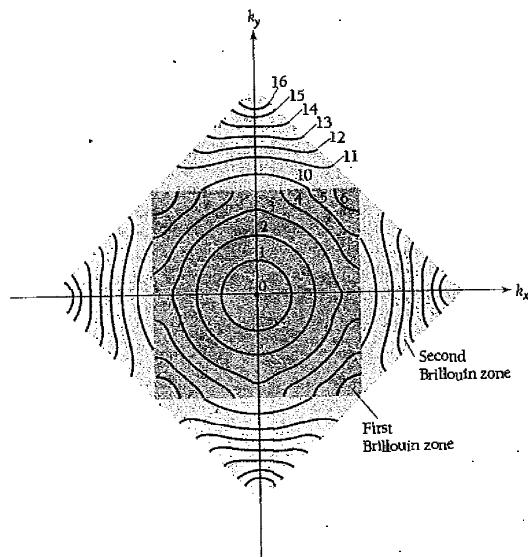


Figure 10.43 Energy contours in electronwells in the first and second Brillouin zones of a hypothetical square lattice.

depends on  $k^2$ , the contour lines of constant energy in a two-dimensional  $k$  space are simply circles of constant  $k$ , as in Fig. 10.43, for such  $k$  values.

With increasing  $k$  the constant-energy contour lines become progressively closer together and also more and more distorted. The reason for the first effect is merely that  $E$  varies with  $k^2$ . The reason for the second is almost equally straightforward. The closer an electron is to the boundary of a Brillouin zone in  $k$ -space, the closer it is to being reflected by the actual crystal lattice. But in particle terms the reflection occurs by virtue of the interaction of the electron with the periodic array of positive ions that occupy the lattice points, and the stronger the interaction, the more the electron's energy is affected.

#### Origin of Forbidden Bands

Figure 10.44 shows how  $E$  varies with  $k$  in the  $x$  direction. As  $k$  approaches  $\pi/a$ ,  $E$  increases more slowly than  $\hbar^2 k^2/2m$ , the free-particle figure. At  $k = \pi/a$ ,  $E$  has two values, the lower belonging to the first Brillouin zone and the higher to the second zone. There is a definite gap between the possible energies in the first and second Brillouin zones which corresponds to a forbidden band. The same pattern continues as successively higher Brillouin zones are reached.

The energy discontinuity at the boundary of a Brillouin zone follows from the fact that the limiting values of  $k$  correspond to standing waves rather than traveling

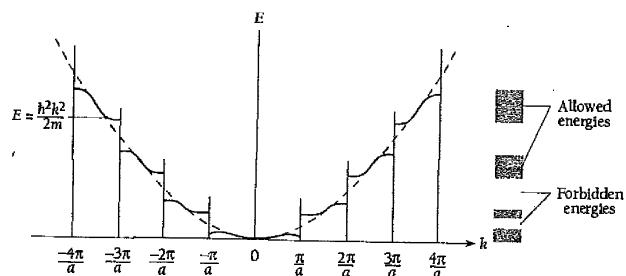


Figure 10.44 Electron energy  $E$  versus wave number  $k$  in the  $k_x$  direction. The dashed line shows how  $E$  varies with  $k$  for a free electron, as given by Eq. (10.22).

waves. For clarity we consider electrons moving in the  $x$  direction; extending the argument to any other direction is straightforward. When  $k = \pm\pi/a$ , as we have seen, the waves are Bragg-reflected back and forth, and so the only solutions of Schrödinger's equation consist of standing waves whose wavelength is equal to the periodicity of the lattice. There are two possibilities for these standing waves for  $n = 1$ , namely,

$$\psi_1 = A \sin \frac{\pi x}{a} \quad (10.23)$$

$$\psi_2 = A \cos \frac{\pi x}{a} \quad (10.24)$$

The probability densities  $|\psi_1|^2$  and  $|\psi_2|^2$  are plotted in Fig. 10.45. Evidently  $|\psi_1|^2$  has its minima at the lattice points occupied by the positive ions, while  $|\psi_2|^2$  has its

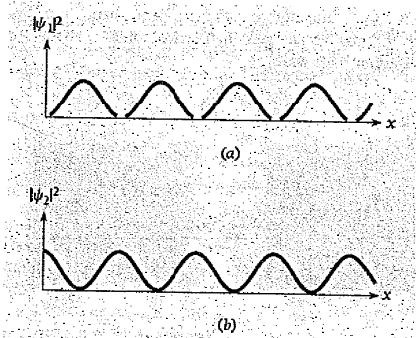


Figure 10.45 Distributions of the probability densities  $|\psi_1|^2$  and  $|\psi_2|^2$ .

### Effective Mass

**Table 10.2** Effective Mass Ratios  $m^*/m$  at the Fermi Surface in Some Metals

Metal		$m^*/m$
Lithium	Li	1.2
Beryllium	Be	1.6
Sodium	Na	1.2
Aluminum	Al	0.97
Cobalt	Co	14
Nickel	Ni	28
Copper	Cu	1.01
Zinc	Zn	0.85
Silver	Ag	0.99
Platinum	Pt	13

Because an electron in a crystal interacts with the crystal lattice, its response to an external electric field is not the same as that of a free electron. Remarkably enough, the most important results of the free-electron theory of metals discussed in Secs. 9.9 and 9.10 can be incorporated in the more realistic band theory merely by replacing the electron mass  $m$  by an average effective mass  $m^*$ . For example, Eq. (9.56) for the Fermi energy is equally valid in the band theory when  $m^*$  is used in place of  $m$ . Table 10.2 is a list of effective mass ratios  $m^*/m$  for several metals.

maxima at the lattice points. Since the charge density corresponding to an electron wave function  $\psi$  is  $-e|\psi|^2$ , the charge density in the case of  $\psi_1$  is concentrated between the positive ions, while in the case of  $\psi_2$ , it is concentrated at the positive ions. The potential energy of an electron in a lattice of positive ions is greatest midway between each pair of ions and least at the ions themselves, so the electron energies  $E_1$  and  $E_2$  associated with the standing waves  $\psi_1$  and  $\psi_2$  are different. No other solutions are possible when  $k = \pm\pi/a$  and accordingly no electron can have an energy between  $E_1$  and  $E_2$ .

Figure 10.46 shows the distribution of electron energies that corresponds to the Brillouin zones pictured in Fig. 10.43. At low energies (in this hypothetical situation for  $E < \sim 2$  eV) the curve is almost exactly the same as that of Fig. 9.11 based on the free-electron theory. This is not surprising since at low energies  $k$  is small and the electrons in a periodic lattice then do behave like free electrons.

With increasing energy, however, the number of available energy states goes beyond that of the free-electron theory owing to the distortion of the energy contours by the lattice. Hence there are more different  $k$  values for each energy. Then, when  $k = \pm\pi/a$ , the energy contours reach the boundaries of the first zone and energies higher than about 4 eV (in this particular model) are forbidden for electrons in the  $k_x$  and  $k_y$  directions although permitted in other directions. As the energy goes farther and farther beyond 4 eV, the available energy states become restricted more and more to the corners of the zone, and  $n(E)$  falls. Finally, at approximately  $6\frac{1}{2}$  eV, there are no more states and  $n(E) = 0$ . The lowest possible energy in the second zone is somewhat less than 10 eV and another curve similar in shape to the first begins. Here the gap between the possible energies in the two zones is about 3 eV, and so the forbidden band is about 3 eV wide.

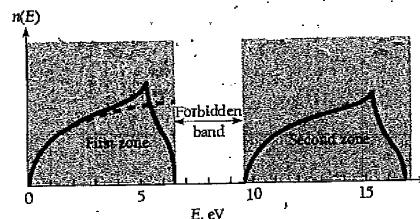


Figure 10.46 The distributions of electron energies in the Brillouin zones of Fig. 10.43. The dashed line is the distribution predicted by the free-electron theory.

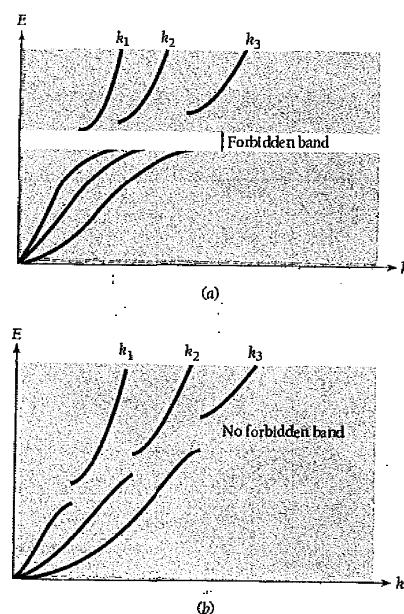


Figure 10.47  $E$  versus  $k$  curves for three directions in two crystals. In (a) there is a forbidden band, in (b) the allowed energy bands overlap and there is no forbidden band.

Although there must be an energy gap between successive Brillouin zones in any given direction, the various gaps may overlap permitted energies in other directions so that there is no forbidden band in the crystal as a whole. Figure 10.47 contains graphs of  $E$  versus  $k$  for three directions (a) in a crystal that has a forbidden band and (b) in a crystal whose allowed bands overlap sufficiently to avoid having a forbidden band.

As we know, the electrical behavior of a solid depends on the degree of occupancy of its energy bands as well as on its band structure. Figure 10.48a shows the first and second Brillouin zones of a hypothetical two-dimensional insulator. The first zone is filled with electrons, and the energy gap between this zone and the second is much wider than  $kT$ . This corresponds to the situation shown in Fig. 10.24 where the insulator is diamond. In Fig. 10.48b the zones are the same, but the first zone is only half filled. This corresponds to the situation shown in Fig. 10.22, and the material is analogous to a metal such as sodium whose atoms have one valence electron each. In Fig. 10.48c the energies in the second zone overlap those in the first zone, so the valence electrons partly occupy both zones. This corresponds to the situation shown in Fig. 10.47b, and the material is analogous to a metal such as magnesium which has two valence electrons per atom.

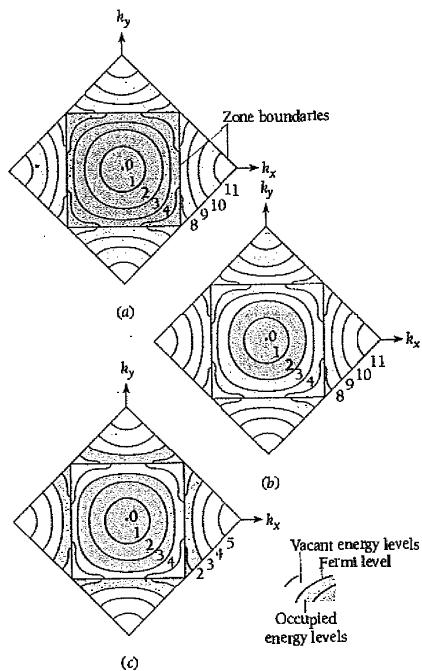


Figure 10.48 Electron energy contours and Fermi levels in three types of solid: (a) insulator; (b) monovalent metal; (c) divalent metal. Energies are in electronvolts.

## 10.9 SUPERCONDUCTIVITY

*No resistance at all, but only at very low temperatures (so far)*

Electrical conductors, even the very best, resist to some extent the flow of charge through them at ordinary temperatures. At very low temperatures, however, most metals, many alloys, and some chemical compounds allow current to pass freely through them. This phenomenon is called **superconductivity**.

Superconductivity was discovered in 1911 by the Dutch physicist Heike Kamerlingh Onnes. He found that, down to 4.15 K, the resistance of a mercury sample decreased with temperature as other metals do (see Fig. 10.18). At  $T_c = 4.15$  K, though, the resistance fell sharply to as close to zero as his instruments could measure (Fig. 10.49). The critical temperature  $T_c$  for other superconducting elements varies from less than 0.1 K to nearly 10 K. As we shall see later, it is significant that elements which are ordinarily good conductors, such as copper and silver, do not become

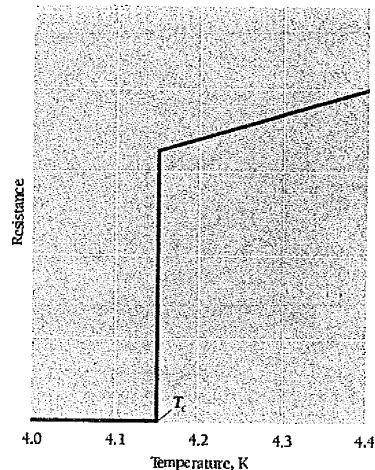


Figure 10.49 Resistance of a mercury sample at low temperature. Below the critical temperature of  $T_c = 4.15$  K mercury is a superconductor with zero electrical resistance.

superconducting when cooled. The highest critical temperatures, as much as 134 K, are found in certain ceramic materials.

Does a superconductor actually have zero resistance or just very little? To find out, currents have been set up in superconducting wire loops and the resulting magnetic fields monitored, sometimes for years. No decrease in such currents has ever been found; superconductors do have no resistance at all.

#### Magnetic Effects

The presence of a magnetic field causes the critical temperature of type I superconductors to decrease in the manner shown in Fig. 10.50. If the magnetic field exceeds a certain critical value  $B_c$ , which depends on the material and its temperature, its superconductivity disappears altogether. Such materials are superconductors only for values of  $T$  and  $B$  below their respective curves and are normal conductors for values of  $T$  and  $B$  above these curves. The critical field  $B_c$  would be a maximum at 0 K.

Table 10.3 gives critical temperatures and critical magnetic fields  $B_c(0)$  extrapolated to 0 K for several type I superconductors. The critical fields are all quite low, less than 0.1 T, so type I superconductors cannot be used for the coils of strong electromagnets.

Superconductors are perfectly diamagnetic—no magnetic field can exist inside them under any circumstances. If we put a sample of a superconductor in a magnetic field weaker than the critical field and then reduce the temperature below  $T_c$ , the field is expelled from the interior of the sample (Fig. 10.51). What happens is that currents appear

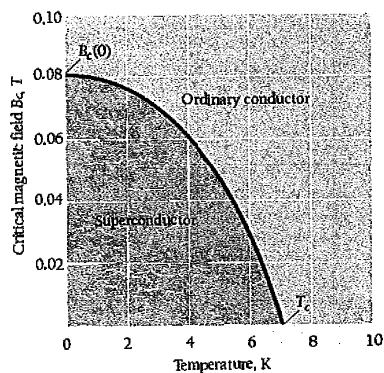


Figure 10.50 Variation of the critical magnetic field  $B_c$  with temperature for lead. Below the curve, lead is a superconductor; above the curve, it is an ordinary conductor.

on the surface of the sample whose magnetic fields exactly cancel the original field inside it. This Meissner effect would not occur in an ordinary conductor whose resistance we can imagine reduced to zero; it is characteristic only of superconductivity, which is evidently a unique state of matter in respects other than ability to conduct electric current.

Type I superconductors exist only in two states, normal and superconducting. Type II superconductors, which were discovered several decades later and are usually alloys, have an intermediate state as well. Such materials have two critical magnetic fields,  $B_{c1}$  and  $B_{c2}$  (Fig. 10.52). For an applied magnetic field less than  $B_{c1}$ , a type II superconductor behaves just like its type I counterpart when  $B < B_c$ ; it is superconducting with no magnetic field in its interior. When  $B > B_{c2}$ , a type II superconductor exhibits normal behavior, again like a type I superconductor. However, in applied fields between  $B_{c1}$  and  $B_{c2}$ , a type II superconductor is in a mixed state in which it

**Table 10.3** Critical Temperatures and Critical Magnetic Fields (at  $T = 0$ ) of Some Type I Superconductors

Superconductor	$T_c$ , K	$B_c(0)$ , T
Al	1.18	0.0105
Hg	4.15	0.0411
In	3.41	0.0281
Pb	7.19	0.0803
Sn	3.72	0.0305
Zn	0.85	0.0054

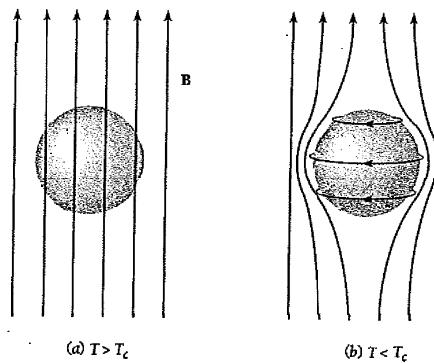


Figure 10.51 The Meissner effect. (a) An applied magnetic field can exist inside a superconductor at temperatures above its critical temperature  $T_c$ . (b) When the superconductor is then cooled below  $T_c$ , surface currents appear whose effect is to expel the magnetic field from the interior of the superconductor.

contains some magnetic flux but is superconducting. The stronger the external field, the more flux penetrates the material, up to the higher critical field  $B_{c2}$ .

A type II superconductor behaves as though it consists of filaments of normal and of superconducting matter mixed together. A magnetic field can exist in the normal filaments, while the superconducting filaments are diamagnetic and resistanceless like

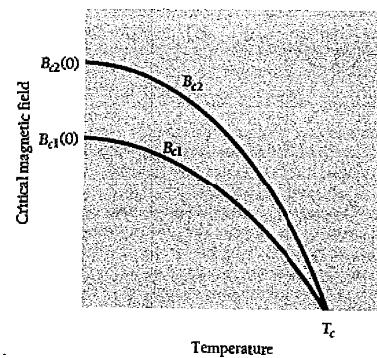


Figure 10.52 Variation of the critical magnetic fields  $B_{c1}$  and  $B_{c2}$  with temperature for a type II superconductor. For magnetic fields between  $B_{c1}$  and  $B_{c2}$  the material is in a mixed state in which it is superconducting but a magnetic field can exist in its interior.

**Table 10.4 Critical Temperatures and Upper Critical Magnetic Fields (at  $T = 0$ ) of Some Type II Superconductors**

Superconductor	$T_c$ , K	$B_{c2}(0)$ , T
$\text{Nb}_3\text{Sn}$	18.0	24.5
$\text{Nb}_3\text{Ge}$	23.2	38
$\text{Nb}_3\text{Al}$	18.7	32.4
$\text{Nb}_3(\text{AlGe})$	20.7	44
$\text{V}_3\text{Ge}$	14.8	2.08
$\text{V}_3\text{Si}$	16.9	2.35
$\text{PbMoS}$	14.4	6.0

type I superconductors. Because  $B_{c2}$  can be quite high (Table 10.4), type II superconductors are used to make high-field (up to 20 T) magnets for particle accelerators, fusion reactors, magnetic resonance imagery, and experimental maglev (magnetic levitation) trains in which magnetic forces provide both propulsion and frictionless support.

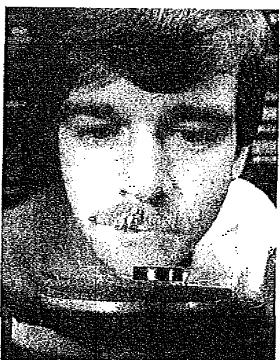
### High-Temperature Superconductors

Despite much effort, until 1986 no superconductor was known whose critical temperature was higher than 27 K. In that year Alex Müller and Georg Bednorz, working in Switzerland, studied a class of ceramic materials that had never before been suspected of superconducting behavior. They discovered an oxide of lanthanum, barium, and copper for which  $T_c$  was 30 K, and soon afterward others extended their approach to produce superconductors with critical temperatures of as high as 134 K ( $-139^\circ\text{C}$ ) for an oxide of mercury, barium, calcium, and copper. (This material has an even higher critical temperature when under pressure.) Although still extremely cold by everyday standards, such temperatures are above the 77-K boiling point of liquid nitrogen, which is cheap (cheaper than milk) and readily available, unlike the liquid helium needed for earlier superconductors.

The new superconductors are all type II and some have high  $B_{c2}$  values. The ceramic crystals consist of layers of copper oxide sandwiched between layers of the other metal oxides. The superconducting occurs in the copper oxide, normally an insulator. Despite much study, the exact mechanism of current flow remains unknown, but it is definitely not the same as in ordinary superconductors.

A number of problems have prevented the wide use of the new superconductors thus far. For instance, like other ceramics crystals they are brittle and difficult to make into wires, cannot carry high currents, and tend to be unstable over long periods. However, methods have been devised to overcome or sidestep these difficulties; one is to encase granules of superconducting material in silver tubes that are then drawn into thin filaments and finally bundled into cables or ribbons. For electric power transmission, the superconducting cables are placed in an insulated pipe through which liquid nitrogen is circulated. The result is not necessarily cheaper than a copper cable that can carry the same current but it is much smaller and lighter. This makes superconducting pipes attractive in such applications as adding electric distribution capacity by replacing copper cables in places where cable ducts are already full, a common situation in cities.

A material that is superconducting at room temperature would revolutionize technology. In addition, by reducing the waste of electrical energy (about 10 percent of the electrical energy generated in the United States is lost as heat in transmission lines), the rate at which the world's resources are being depleted would be reduced. Since 1986 such a material no longer seems inconceivable.



Magnetic levitation. A small permanent magnet is floating freely above a high-temperature superconductor cooled with liquid nitrogen. The magnetic field of the magnet induces electric currents in the superconductor which lead to a zero resultant field inside the superconductor. The magnetic field of these currents outside the superconductor repels the magnet.

## 10.10 BOUND ELECTRON PAIRS

### *The key to superconductivity*

The origin of superconductivity remained a mystery until the Bardeen-Cooper-Schrieffer (BCS) theory of 1957. An earlier hint of the direction such a theory should take was the discovery that the critical temperatures  $T_c$  of the isotopes of a superconducting element decrease with increasing atomic mass. For instance, in mercury  $T_c$  is 4.161 K in  $^{199}\text{Hg}$  but only 4.126 K in  $^{201}\text{Hg}$ . This isotope effect suggests that the current-carrying electrons in a superconductor do not move independently of the ion lattice (as we might think when we recall that the resistance of ordinary conductors arises from the scattering of these electrons by lattice defects and vibrations) but instead are somehow interacting with the lattice.

The nature of the interaction became clear when Leon Cooper showed how two electrons in a superconductor could form a bound state despite their coulomb repulsion. What happens is that the lattice is slightly deformed as an electron moves through it, with the positive ions in the electron's path being displaced toward it. The deformation produces a region of increased positive charge. Another electron moving through this polarized region will be attracted by the greater concentration of positive charge there. If the attraction is stronger than the repulsion between the electrons, the electrons are effectively coupled together into a Cooper pair with the deformed lattice as the intermediary.

The electron-lattice-electron interaction does not keep the electrons a fixed distance apart. In fact, the theory shows that they must be moving in opposite directions, and their correlations may persist over lengths as great as  $10^{-6}$  m. The binding energy of



John Bardeen (1908–1991) was born in Madison, Wisconsin, and studied electrical engineering at the University of Wisconsin and solid-state physics at Princeton University. After working at several universities and, during World War II, at the Naval Ordnance Laboratory, he went to Bell Telephone Laboratories in 1945 where he joined a semiconductor research group led by William Shockley. In 1948 the group produced the first transistor, for which Shockley, Bardeen, and their collaborator

Walter Brattain received a Nobel Prize in 1956. Bardeen later said, "I knew the transistor was important, but I never foresaw the revolution in electronics it would bring."

In 1951 Bardeen left Bell Labs for the University of Illinois where, together with Leon Cooper and J. Robert Schrieffer, he developed the theory of superconductivity. Compared with his earlier work on the transistor, "Superconductivity was more difficult to solve, and it required some radically new concepts." According to the theory, the motions of two electrons can become correlated through their interactions with a crystal lattice, which enables the pair to move with complete freedom through the crystal. Bardeen received his second Nobel Prize in 1972 for this theory along with Cooper and Schrieffer; he was the first person to receive two such prizes in the same field.

a Cooper pair, called the energy gap  $E_g$ , is of the order of  $10^{-3}$  eV, which is why superconductivity is a low-temperature phenomenon. The energy gap can be measured by directing microwave radiation of frequency  $\nu$  at a superconductor. When  $h\nu \geq E_g$ , strong absorption occurs as the Cooper pairs break apart.

The BCS theory relates the energy gap of a superconductor at 0 K to its critical temperature  $T_c$  by the formula

$$\text{Energy gap at } 0 \text{ K} \quad E_g(0) = 3.53kT_c \quad (10.25)$$

Equation (10.25) agrees fairly well with the observed values of  $E_g$  and  $T_c$ . At temperatures above 0 K, some Cooper pairs break up. The resulting individual electrons interact with the remaining Cooper pairs and reduce the energy gap (Fig. 10.53). Finally,

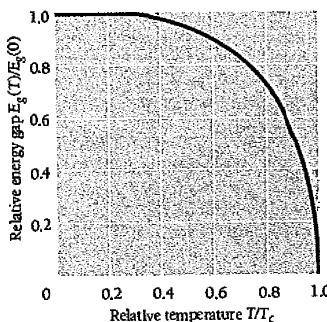


Figure 10.53 Variation of the superconducting energy gap with temperature. Here  $E_g(T)$  is the energy gap at the temperature  $T$  and  $E_g(0)$  is the gap at  $T = 0$ ;  $T_c$  is the critical temperature of the material.

at the critical temperature  $T_c$ , the energy gap disappears, there are no more Cooper pairs, and the material is no longer superconducting.

The electrons in a Cooper pair have opposite spins, so the pair has a total spin of zero. As a result, the electron pairs in a superconductor are bosons (unlike individual electrons, which have spins of  $\frac{1}{2}$  and are fermions), and any number of them can exist in the same quantum state at the same time. When there is no current in the superconductor, the linear momenta of the electrons in a Cooper pair are equal and opposite for a total of zero. All the pairs are then in the same ground state and make up a giant system the size of the superconductor. A single wave function represents this system, whose total energy is less than that of a system of the same number of electrons with a Fermi energy distribution.

A current in a superconductor involves the entire system of electron pairs acting as a unit. Every pair now has a non-zero momentum. To alter such a current means that the correlated states of motion of *all* the electron pairs, not just the states of motion of some individual electrons as in an ordinary conductor, must be changed. Because such a change requires a relatively large amount of energy, the current persists indefinitely.

### Flux Quantization

Figure 10.54 shows a superconducting ring of area  $A$  that carries a current. The amount of magnetic flux  $\Phi = BA$  passes through the ring as a result. According to Faraday's law of electromagnetic induction, any change in the flux will change the current in the ring so as to oppose the change in flux. Because the ring has no resistance, the change in flux will be perfectly canceled out. The flux  $\Phi$  therefore is permanently trapped.

Because the phase of the wave function of the Cooper pairs in the ring must be continuous around the ring, it turns out that  $\Phi$  is quantized. The only values that  $\Phi$  can have are

$$\text{Flux quantization} \quad \Phi = n \left( \frac{\hbar}{2e} \right) = n\Phi_0 \quad n = 1, 2, 3, \dots \quad (10.26)$$

The quantum of magnetic flux is

$$\text{Flux quantum} \quad \Phi_0 = \frac{\hbar}{2e} = 2.068 \times 10^{-15} \text{ T} \cdot \text{m}^2$$

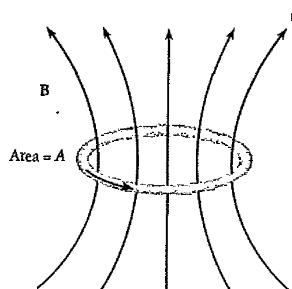


Figure 10.54 The magnetic flux  $\Phi = BA$  that passes through a superconducting ring can only have the values  $\Phi = n\Phi_0$  where  $\Phi_0$  is the flux quantum and  $n = 1, 2, 3, \dots$

if undisturbed, and the electron scattering that leads to resistance in an ordinary conductor does not occur.

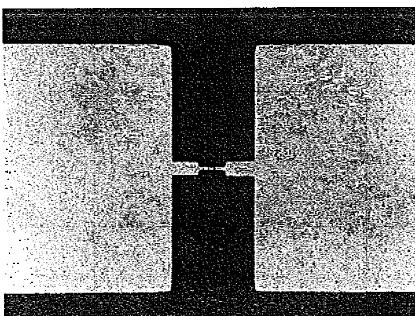
A material with large-amplitude lattice vibrations may be only a fair conductor at ordinary temperatures because electron scattering takes place frequently. However, the same ease of lattice deformation means more strongly bound Cooper pairs at low temperatures, and hence the material is more likely to be a superconductor then. Good conductors, such as copper and silver, have small lattice vibrations at ordinary temperatures, which means their lattices are unable to mediate the formation of Cooper pairs at low temperatures and so they do not become superconducting. Such metals as mercury, tin, and lead have large lattice vibrations at ordinary temperatures and so are poorer conductors than copper and silver, but they are superconductors at low temperatures.

### Josephson Junctions

As we learned in Chap. 5, the wave nature of a moving particle allows it to tunnel through a barrier that, in classical physics, it could not penetrate. Thus a small but detectable current of electrons can tunnel through a thin insulating layer between two metals. In 1962 Brian Josephson, then a graduate student at Cambridge University, predicted that Cooper pairs could tunnel through what is now called a Josephson junction, a thin insulating layer between two superconductors. The wave functions of the Cooper pairs on each side of the junction penetrate the insulating layer with exponentially decreasing amplitudes, just as the wave functions of individual electrons would. If the layer is thin enough, less than 2 nm in practice, the wave functions overlap sufficiently to become coupled together, and the Cooper pairs they describe can then pass through the junction. Josephson shared the 1975 Nobel Prize in physics for his work.

In the dc Josephson effect, the current through a Josephson junction that has no voltage across it is given by

$$\text{dc Josephson effect} \quad I_J = I_{\max} \sin \phi \quad (10.27)$$



The small rectangle at the center of this photograph is a Josephson junction 1.25  $\mu\text{m}$  wide.

Here  $\phi$  is the phase difference between the wave functions of the Cooper pairs on either side of the junction. The value  $I_{\max}$  of the maximum junction current depends on the thickness of the insulating layer and is quite small, between 1  $\mu\text{A}$  and 1 mA in a Nb-NbO-Nb junction, for example.

When a voltage  $V$  is applied across a Josephson junction, the phase difference  $\phi$  increases with time at the rate

$$\text{ac Josephson effect} \quad \nu = \frac{d\phi}{dt} = \frac{2V}{h} \quad (10.28)$$

As a result,  $I_J$  varies sinusoidally with time, which constitutes the ac Josephson effect. The value of  $2e/h$  is 483.5979 THz/volt. Because  $\nu$  is proportional to  $V$  and can be measured accurately, for instance by finding the frequency of the em radiation emitted by the junction, the ac Josephson effect enables very precise voltage determinations to be made. In fact, the effect is the basis for the present definition of the volt: one volt is the potential difference across a Josephson junction that produces oscillations at a frequency of 483,5979 THz.

Josephson junctions are used in extremely sensitive magnetometers called SQUIDS—superconducting quantum interference devices. SQUIDs vary in detail, but all make use of the fact that the maximum current in a superconducting ring that contains a Josephson junction varies periodically as the magnetic flux through the ring changes. The periodicity is interpreted as an interference effect involving the wave functions of the Cooper pairs. Magnetic field changes as small as  $10^{-21}$  T can be detected by SQUIDs, which among other applications permits sensing the weak magnetic fields produced by biological currents such as those in the brain.

## EXERCISES

I pass with relief from the tossing sea of Cause and Theory to the firm ground of Result and Fact. —Winston Churchill

### 10.2 Ionic Crystals

1. The ion spacings and melting points of the sodium halides are as follows:

	NaF	NaCl	NaBr	NaI
Ion spacing, nm	0.23	0.28	0.29	0.32
Melting point, °C	988	801	740	660

Explain the regular variation in these quantities with halogen atomic number.

2. Show that the first five terms in the series for the Madelung constant of NaCl are

$$\alpha = 6 - \frac{12}{\sqrt{2}} + \frac{8}{\sqrt{3}} - \frac{6}{2} + \frac{24}{\sqrt{5}} - \dots$$

3. (a) The ionization energy of potassium is 4.34 eV and the electron affinity of chlorine is 3.61 eV. The Madelung constant for the KCl structure is 1.748 and the distance between ions of opposite sign is 0.314 nm. On the basis of these data only,

compute the cohesive energy of KCl. (b) The observed cohesive energy of KCl is 6.42 eV per ion pair. On the assumption that the difference between this figure and that obtained in a is due to the exclusion-principle repulsion, find the exponent  $n$  in the formula  $Br^{-n}$  for the potential energy arising from this source.

4. Repeat Exercise 3 for LiCl, in which the Madelung constant is 1.748, the ion spacing is 0.257 nm, and the observed cohesive energy is 6.8 eV per ion pair. The ionization energy of Li is 5.4 eV.

### 10.4 Van der Waals Bond

5. The Joule-Thomson effect refers to the drop in temperature a gas undergoes when it passes slowly from a full container to an empty one through a porous plug. Since the expansion is into a rigid container, no mechanical work is done. Explain the Joule-Thomson effect in terms of the van der Waals attraction between molecules.
6. Van der Waals forces can hold inert gas atoms together to form solids at low temperatures, but they cannot hold such atoms together to form molecules in the gaseous state. Why not?

7. What is the effect on the cohesive energy of ionic and covalent crystals of (a) van der Waals forces and (b) zero-point oscillations of the ions and atoms about their equilibrium positions?

#### 10.5 Metallic Bond

8. Lithium atoms, like hydrogen atoms, have only a single electron in their outer shells, yet lithium atoms do not join together to form  $\text{Li}_2$  molecules the way hydrogen atoms form  $\text{H}_2$  molecules. Instead, lithium is a metal with each atom part of a crystal lattice. Why?
9. Does the "gas" of freely moving electrons in a metal include all the electrons present? If not, which electrons are members of the "gas"?
10. Gold has an atomic mass of 197 u, a density of  $19.3 \times 10^3 \text{ kg/m}^3$ , a Fermi energy of 5.54 eV, and a resistivity of  $2.04 \times 10^{-8} \Omega \cdot \text{m}$ . Estimate the mean free path in atom spacings between collisions of the free electrons in gold under the assumption that each gold atom contributes one electron to the electron gas.
11. Silver has an atomic mass of 108 u, a density of  $10.5 \times 10^3 \text{ kg/m}^3$ , and a Fermi energy of 5.51 eV. On the assumptions that each silver atom contributes one electron to the electron gas and that the mean free path of the electrons is 200 atom spacings, estimate the resistivity of silver. (The actual resistivity of silver at 20°C is  $1.6 \times 10^{-8} \Omega \cdot \text{m}$ .)

#### 10.6 Band Theory of Solids

12. What is the basic physical principle responsible for the presence of energy bands rather than specific energy levels in a solid?
13. How are the band structures of insulators and semiconductors similar? How are they different?
14. What are the two combinations of band structure and occupancy by electrons that can cause a solid to be a metal?
15. (a) Why are some solids transparent to visible light and others opaque? (b) The forbidden band is 1.1 eV in silicon and 6 eV in diamond. To what wavelengths of light are these substances transparent?
16. The forbidden band is 0.7 eV in germanium and 1.1 eV in silicon. How does the conductivity of germanium compare with that of silicon at (a) very low temperatures and (b) room temperature?
17. (a) When germanium is doped with aluminum, is the result an *n*-type or a *p*-type semiconductor? (b) Why?

#### 10.8 Energy Bands: Alternative Analysis

18. Compare the de Broglie wavelength of an electron in copper with the 7.04-eV Fermi energy with the 0.256-nm spacing of the copper atoms.
19. Draw the third Brillouin zone of the two-dimensional square lattice whose first two Brillouin zones are shown in Fig. 10.41.
20. Find the ratio between the kinetic energies of an electron in a two-dimensional square lattice which has  $k_x = k_y = \pi/a$  and an electron which has  $k_x \approx \pi/a$ ,  $k_y = 0$ .
21. Phosphorus is present in germanium sample. Assume that one of its five valence electrons revolves in a Bohr orbit around each  $\text{P}^{3+}$  ion in the germanium lattice. (a) If the effective mass of the electron is  $0.17 m_e$  and the dielectric constant of germanium is 16, find the radius of the first Bohr orbit of the electron. (b) The energy gap between the valence and conduction bands in germanium is 0.65 eV. How does the ionization energy of the above electron compare with this energy and with  $kT$  at room temperature?
22. Repeat Exercise 21 for a silicon sample that contains arsenic. The effective mass of an electron in silicon is about  $0.31 m_e$ , the dielectric constant of silicon is 12, and the energy gap in silicon is 1.1 eV.
23. The effective mass  $m^*$  of a current carrier in a semiconductor can be directly determined by means of a cyclotron resonance experiment in which the carriers (whether electrons or holes) move in helical orbits about the direction of an externally applied magnetic field  $B$ . An alternating electric field is applied perpendicular to  $B$ , and resonant absorption of energy from this field occurs when its frequency  $\nu$  is equal to the frequency of revolution  $\nu_c$  of the carrier. (a) Derive an equation for  $\nu_c$  in terms of  $m^*$ ,  $e$ , and  $B$ . (b) In a certain experiment,  $B = 0.1 \text{ T}$  and maximum absorption is found to occur at  $\nu = 1.4 \times 10^{10} \text{ Hz}$ . Find  $m^*$ . (c) Find the maximum orbital radius of a charge carrier in this experiment whose speed is  $3 \times 10^5 \text{ m/s}$ .

#### 10.9 Superconductivity

#### 10.10 Bound Electron Pairs

24. The actual energy gap at 0 K in lead is  $2.73 \times 10^{-3} \text{ eV}$ . (a) What is the prediction of the BCS theory for this energy gap? (b) Radiation of what minimum frequency could break apart Cooper pairs in lead at 0 K? In what part of the em spectrum is such radiation?
25. A voltage of  $5.0 \mu\text{V}$  is applied across a Josephson junction. What is the frequency of the radiation emitted by the junction?
26. A SQUID magnetometer that uses a superconducting ring 2.0 mm in diameter indicates a change in the magnetic flux through it of 5 flux quanta. What is the corresponding magnetic field change?