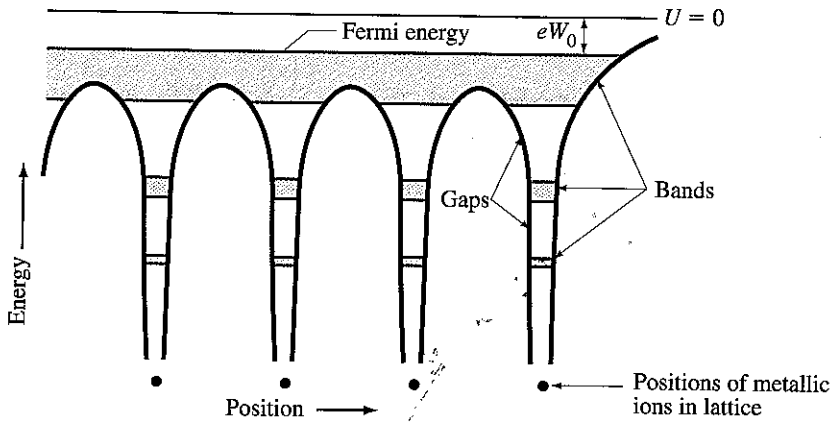

FIGURE 40–29 Potential energy for an electron in a metal crystal, with deep potential wells in the vicinity of each ion in the crystal lattice.

The simple model of an electron gas presented here provides good explanations for the electrical and thermal properties of conductors. But it does not explain why some materials are good conductors and others are good insulators. To provide an explanation, our model of electrons inside a metal moving in a uniform potential well needs to be refined to include the effect of the lattice. Figure 40–29 shows a "periodic" potential that takes into account the attraction of electrons for each atomic ion in the lattice. Here we have taken $U = 0$ for an electron free of the metal; so within the metal, electron energies are less than zero (just as for molecules, or for the H atom in which the ground state has $E = -13.6\,\text{eV}$). The quantity eW_0 represents the minimum energy to remove an electron from the metal, where W_0 is the *work function* (see Section 37–2). The crucial outcome of putting a periodic potential (more easily approximated with narrow square wells) into the Schrödinger equation is that the allowed energy states are divided into *bands*, with energy gaps in between. Only electrons in the highest band, close to the Fermi level, are able to move about freely within the metal crystal. In the next Section we will see physically why there are bands and how they explain the properties of conductors, insulators, and semiconductors.

40–7 Band Theory of Solids

We saw in Section 40–1 that when two hydrogen atoms approach each other, the wave functions overlap, and the two 1*s* states (one for each atom) divide into two states of different energy. (As we saw, only one of these states, $S = 0$, has low enough energy to give a bound H_2 molecule.) Figure 40–30a shows this situation for 1*s* and 2*s* states for two atoms: as the two atoms get closer (toward the left in Fig. 40–30a), the 1*s* and 2*s* states split into two levels. If six atoms come together, as in Fig. 40–30b, each of the states splits into six levels. If a large number of atoms come together to form a solid, then each of the original atomic levels becomes a **band** as shown in Fig. 40–30c. The energy levels are so close together in each band that they seem essentially continuous. This is why the spectrum of heated solids (Section 37–1) appears continuous. (See also Fig. 40–15 and its discussion at the start of Section 40–4.)

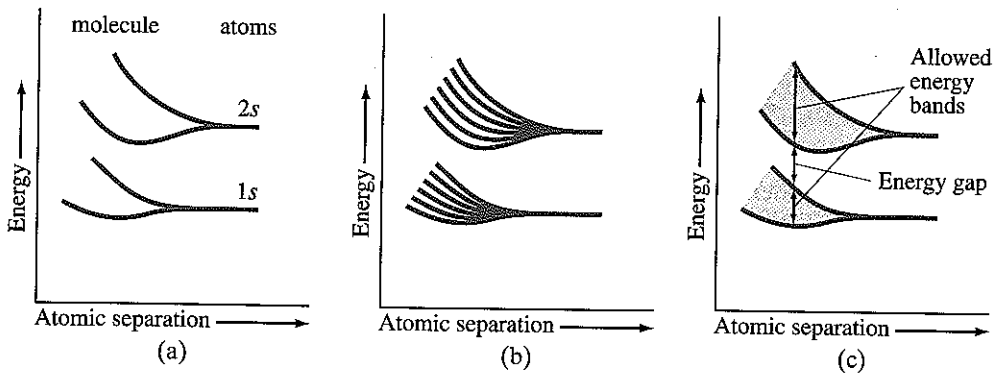

FIGURE 40–30 The splitting of 1*s* and 2*s* atomic energy levels as (a) two atoms approach each other (the atomic separation decreases toward the left on the graph), (b) the same for six atoms, and (c) for many atoms when they come together to form a solid.

The crucial aspect of a good **conductor** is that the highest energy band containing electrons is only partially filled. Consider sodium metal, for example, whose energy bands are shown in Fig. 40–31. The 1*s*, 2*s*, and 2*p* bands are full (just as in a sodium atom) and don't concern us. The 3*s* band, however, is only half full. To see why, recall that the exclusion principle stipulates that in an atom, only two electrons can be in the 3*s* state, one with spin up and one with spin down. These two states have slightly different energy. For a solid consisting of N atoms, the 3*s* band will contain $2N$ possible energy states. A sodium atom has a single 3*s* electron, so in a sample of sodium metal containing N atoms, there are N electrons in the 3*s* band, and N unoccupied states. When a potential difference is applied across the metal, electrons can respond by accelerating and increasing their energy, since there are plenty of unoccupied states of slightly higher energy available. Hence, a current flows readily and sodium is a good conductor. The characteristic of all good conductors is that the highest energy band is only partially filled, or two bands overlap so that unoccupied states are available. An example of the latter is magnesium, which has two 3*s* electrons, so its 3*s* band is filled. But the unfilled 3*p* band overlaps the 3*s* band in energy, so there are lots of available states for the electrons to move into. Thus magnesium, too, is a good conductor.

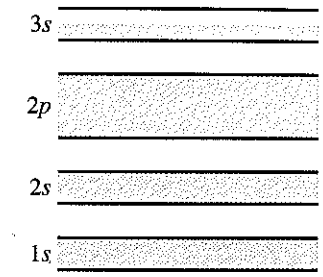

FIGURE 40–31 Energy bands for sodium (Na).

In a material that is a good **insulator**, on the other hand, the highest band containing electrons, called the **valence band**, is completely filled. The next highest energy band, called the **conduction band**, is separated from the valence band by a "forbidden" **energy gap** (or **band gap**), E_g, of typically 5 to 10 eV. So at room temperature (300 K), where thermal energies (that is, average kinetic energy—see Chapter 18) are on the order of $\frac{3}{2}kT \approx 0.04$ eV, almost no electrons can acquire the 5 eV needed to reach the conduction band. When a potential difference is applied across the material, no available states are accessible to the electrons, and no current flows. Hence, the material is a good insulator.

Figure 40–32 compares the relevant energy bands (a) for conductors, (b) for insulators, and also (c) for the important class of materials known as **semiconductors**. The bands for a pure (or **intrinsic**) semiconductor, such as silicon or germanium, are like those for an insulator, except that the unfilled conduction band is separated from the filled valence band by a much smaller energy gap, E_g, typically on the order of 1 eV. At room temperature, a few electrons can acquire enough thermal energy to reach the conduction band, and so a very small current may flow when a voltage is applied. At higher temperatures, more electrons have enough energy to jump the gap. Often this effect can more than offset the effects of more frequent collisions due to increased disorder at higher temperature, so the resistivity of semiconductors can *decrease* with increasing temperature (see Table 25–1). But this is not the whole story of semiconductor conduction. When a potential difference is applied to a semiconductor, the few electrons in the conduction band move toward the positive electrode. Electrons in the valence band try to do the same thing, and a few can because there are a small number of unoccupied states which were left empty by the electrons reaching the conduction band. Such unfilled electron states are called **holes**. Each electron in the valence band that fills a hole in this way as it moves toward the positive electrode leaves behind its own hole, so the holes migrate toward the negative electrode. As the electrons tend to accumulate at one side of the material, the holes tend to accumulate on the opposite side. We will look at this phenomenon in more detail in the next Section.

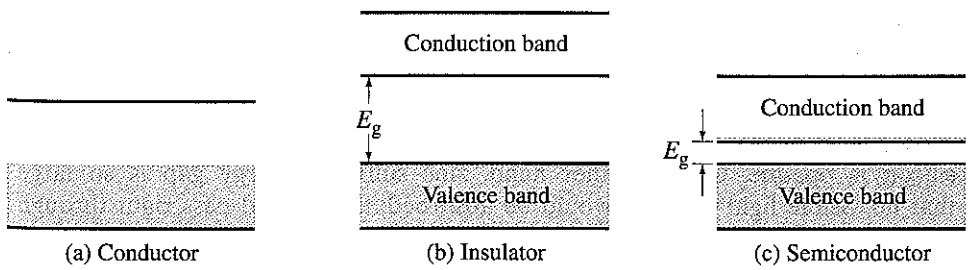

FIGURE 40–32 Energy bands for (a) a conductor, (b) an insulator, which has a large energy gap E_g, and (c) a semiconductor, which has a small energy gap E_g. Shading represents occupied states. Pale shading in (c) represents electrons that can pass from the top of the valence band to the bottom of the conduction band due to thermal agitation at room temperature (exaggerated).

EXAMPLE 40–12 **Calculating the energy gap.** It is found that the conductivity of a certain semiconductor increases when light of wavelength 345 nm or shorter strikes it, suggesting that electrons are being promoted from the valence band to the conduction band. What is the energy gap, E_g, for this semiconductor?

APPROACH The longest wavelength (lowest energy) photon to cause an increase in conductivity has $\lambda = 345\,\text{nm}$, and its energy ($= hf$) equals the energy gap.

SOLUTION The gap energy equals the energy of a $\lambda = 345$-nm photon:

$$E_g = hf = \frac{hc}{\lambda} = \frac{(6.63 \times 10^{-34}\,\text{J}\cdot\text{s})(3.00 \times 10^{8}\,\text{m/s})}{(345 \times 10^{-9}\,\text{m})(1.60 \times 10^{-19}\,\text{J/eV})} = 3.6\,\text{eV}.$$

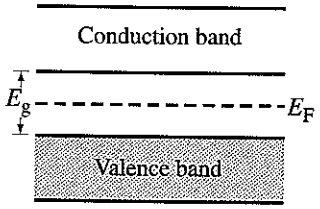

FIGURE 40–33 The Fermi energy is midway between the valence band and the conduction band.

EXAMPLE 40–13 **ESTIMATE** **Free electrons in semiconductors and insulators.** Use the Fermi–Dirac probability function, Eq. 40–14, to estimate the order of magnitude of the numbers of free electrons in the conduction band of a solid containing 10^{21} atoms, assuming the solid is at room temperature ($T = 300\,\text{K}$) and is (*a*) a semiconductor with $E_g \approx 1.1\,\text{eV}$, (*b*) an insulator with $E_g \approx 5\,\text{eV}$. Compare to a conductor.

APPROACH At $T = 0$, all states above the Fermi energy E_F are empty, and all those below are full. So for semiconductors and insulators we can take E_F to be about midway between the valence and conduction bands, Fig. 40–33, and it does not change significantly as we go to room temperature. We can thus use Eq. 40–14 to find the fraction of electrons in the conduction band at room temperature for the two cases.

SOLUTION (*a*) For the semiconductor, the gap $E_g \approx 1.1\,\text{eV}$, so $E - E_F \approx 0.55\,\text{eV}$ for the lowest states in the conduction band. Since at room temperature we have $kT \approx 0.026\,\text{eV}$, then $(E - E_F)/kT \approx 0.55\,\text{eV}/0.026\,\text{eV} \approx 21$ and

$$f(E) = \frac{1}{e^{(E-E_F)/kT} + 1} \approx \frac{1}{e^{21}} \approx 10^{-9}.$$

Thus about 1 atom in 10^9 can contribute an electron to the conductivity.
(*b*) For the insulator with $E - E_F \approx 5.0\,\text{eV} - \frac{1}{2}(5.0\,\text{eV}) = 2.5\,\text{eV}$, we get

$$f(E) \approx \frac{1}{e^{2.5/0.026} + 1} \approx \frac{1}{e^{96}} \approx 10^{-42}.$$

Thus in an ordinary sample containing 10^{21} atoms, there would be no free electrons in an insulator $(10^{21} \times 10^{-42} = 10^{-21})$, about 10^{12} $(10^{21} \times 10^{-9})$ free electrons in a semiconductor, and about 10^{21} free electrons in a good conductor.

PHYSICS APPLIED
Transparency

CONCEPTUAL EXAMPLE 40–14 **Which is transparent?** The energy gap for silicon is 1.14 eV at room temperature, whereas that of zinc sulfide (ZnS) is 3.6 eV. Which one of these is opaque to visible light, and which is transparent?

RESPONSE Visible light photons span energies from roughly 1.8 eV to 3.1 eV ($E = hf = hc/\lambda$ where $\lambda = 400\,\text{nm}$ to 700 nm and $1\,\text{eV} = 1.6 \times 10^{-19}\,\text{J}$). Light is absorbed by the electrons in a material. Silicon's energy gap is small enough to absorb these photons, thus bumping electrons well up into the conduction band, so silicon is opaque. On the other hand, zinc sulfide's energy gap is so large that no visible light photons would be absorbed; they would pass right through the material which would thus be transparent.

40–8 Semiconductors and Doping

Nearly all electronic devices today use semiconductors. The most common are silicon (Si) and germanium (Ge). An atom of silicon or germanium has four outer electrons that act to hold the atoms in the regular lattice structure of the crystal, shown schematically in Fig. 40–34a. Germanium and silicon acquire properties useful for electronics when a tiny amount of impurity is introduced into the crystal structure (perhaps 1 part in 10^6 or 10^7). This is called **doping** the semiconductor. Two kinds of doped semiconductor can be made, depending on the type of impurity used. If the impurity is an element whose atoms have five outer electrons, such as arsenic, we have the situation shown in Fig. 40–34b, with the arsenic atoms holding positions in the crystal lattice where normally silicon atoms would be. Only four of arsenic's electrons fit into the bonding structure. The fifth does not fit in and can move relatively freely, somewhat like the electrons in a conductor. Because of this small number of extra electrons, a doped semiconductor becomes slightly conducting.

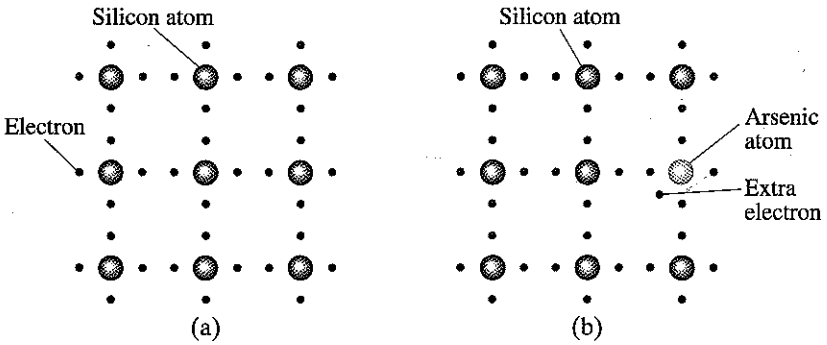

FIGURE 40–34 Two-dimensional representation of a silicon crystal. (a) Four (outer) electrons surround each silicon atom. (b) Silicon crystal doped with a small percentage of arsenic atoms: the extra electron doesn't fit into the crystal lattice and so is free to move about. This is an *n*-type semiconductor.

The density of conduction electrons in an intrinsic (undoped) semiconductor at room temperature is very low, usually less than 1 per 10^9 atoms. With an impurity concentration of 1 in 10^6 or 10^7 when doped, the conductivity will be much higher and it can be controlled with great precision. An arsenic-doped silicon crystal is called an ***n*-type semiconductor** because *negative* charges (electrons) carry the electric current.

In a ***p*-type semiconductor**, a small percentage of semiconductor atoms are replaced by atoms with three outer electrons—such as gallium. As shown in Fig. 40–35a, there is a "hole" in the lattice structure near a gallium atom since it has only three outer electrons. Electrons from nearby silicon atoms can jump into this hole and fill it. But this leaves a hole where that electron had previously been, Fig. 40–35b. The vast majority of atoms are silicon, so holes are almost always next to a silicon atom. Since silicon atoms require four outer electrons to be neutral, this means that there is a net positive charge at the hole. Whenever an electron moves to fill a hole, the positive hole is then at the previous position of that electron. Another electron can then fill this hole, and the hole thus moves to a new location; and so on. This type of semiconductor is called *p-type* because it is the positive holes that seem to carry the electric current. Note, however, that both *p-type* and *n-type* semiconductors have *no net charge* on them.

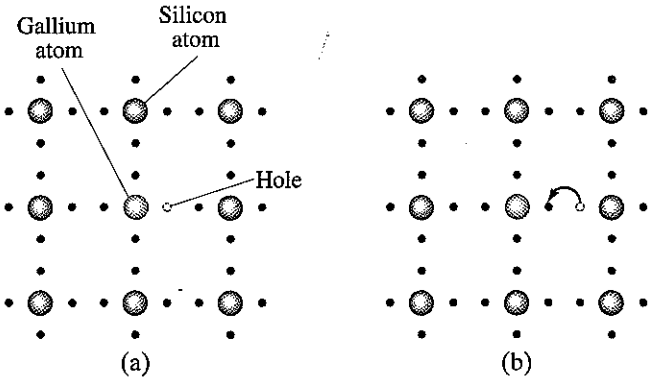

FIGURE 40–35 A *p*-type semiconductor, gallium-doped silicon. (a) Gallium has only three outer electrons, so there is an empty spot, or *hole* in the structure. (b) Electrons from silicon atoms can jump into the hole and fill it. As a result, the hole moves to a new location (to the right in this Figure), to where the electron used to be.

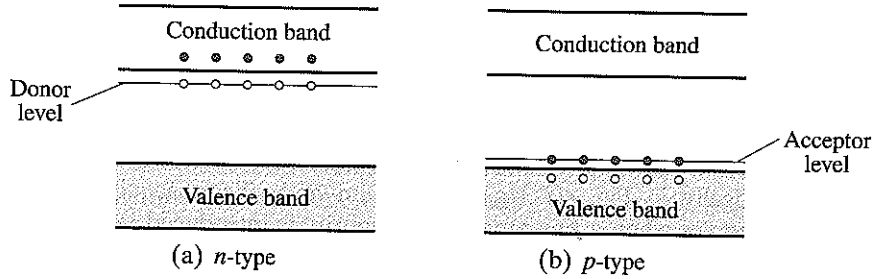

FIGURE 40–36 Impurity energy levels in doped semiconductors.

According to the band theory (Section 40–7), in a doped semiconductor the impurity provides additional energy states between the bands as shown in Fig. 40–36. In an *n*-type semiconductor, the impurity energy level lies just below the conduction band, Fig. 40–36a. Electrons in this energy level need only about 0.05 eV in Si (even less in Ge) to reach the conduction band; this is on the order of the thermal energy, $\frac{3}{2}kT$ (= 0.04 eV at 300 K), so transitions occur readily at room temperature. This energy level can thus supply electrons to the conduction band, so it is called a **donor** level. In *p*-type semiconductors, the impurity energy level is just above the valence band (Fig. 40–36b). It is called an **acceptor** level because electrons from the valence band can easily jump into it. Positive holes are left behind in the valence band, and as other electrons move into these holes, the holes move as discussed earlier.

EXERCISE D Which of the following impurity atoms would produce a *p*-type semiconductor? (*a*) Ge; (*b*) Ne; (*c*) Al; (*d*) As; (*e*) none of the above.

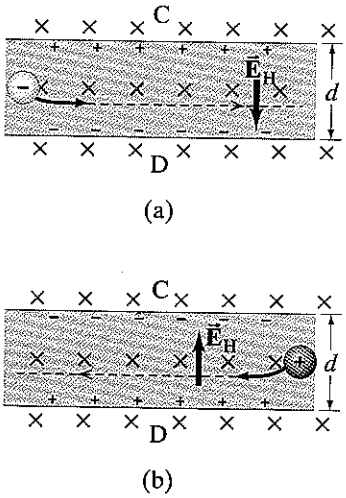

FIGURE 27–32 (Repeated.) The Hall effect. (a) Negative charges moving to the right as the current. (b) The same current, but as positive charges moving to the left.

CONCEPTUAL EXAMPLE 40–15 **Determining charge of conductors.** How can we determine if a *p*-type semiconductor has a current that is really due to the motion of holes? Or, is this just a convenient model?

RESPONSE Recall from Section 27–8 that the Hall effect can be used to distinguish the sign of the charges involved in a current. When placed in a magnetic field, the current in a particular direction can result in a voltage perpendicular to that current due to the magnetic force on the moving charges (Fig. 27–32, repeated here). The direction of this *Hall voltage* depends on the sign of the charges carrying the current. In this way, it has been demonstrated that it really is moving holes that are responsible for the current in a *p*-type semiconductor.

40–9 Semiconductor Diodes

Semiconductor diodes and transistors are essential components of modern electronic devices. The miniaturization achieved today allows many thousands of diodes, transistors, resistors, and so on, to be placed on a single *chip* less than a millimeter on a side. We now discuss, briefly and qualitatively, the operation of diodes and transistors.

When an *n*-type semiconductor is joined to a *p*-type, a ***pn* junction diode** is formed. Separately, the two semiconductors are electrically neutral. When joined, a few electrons near the junction diffuse from the *n*-type into the *p*-type semiconductor, where they fill a few of the holes. The *n*-type is left with a positive charge, and the *p*-type acquires a net negative charge. Thus a potential difference is established, with the *n* side positive relative to the *p* side, and this prevents further diffusion of electrons.