Fig. 14 Energy band structures of (*a*) Si and (*b*) GaAs. Circles (o) indicate holes in the valence bands and dots (•) indicate electrons in the conduction bands.

We note that the general features in Fig. 14 are similar to those in Fig. 13. First of all, the valence bands are simpler than the conduction bands. They are qualitatively similar for most semiconductors because the environments for holes moving in the covalent bonds are similar due to the similar structures in diamond and zincblende. There is a bandgap E_g between the bottom of the conduction band and the top of the valence band. Near the minimum of the conduction band or the maximum of the valence band, the E-p curves are essentially parabolic. For silicon (Fig. 14*a*) the maximum in the valence band occurs at $p = 0$, but the minimum in the conduction band occurs along the [100] direction at $p = p_c$. Therefore, in silicon, when an electron makes a transition from the maximum point in the valence band to the minimum point in the conduction band, not only an energy change ($\geq E_g$) but also some momentum change ($\geq p_c$) is required.

For gallium arsenide (Fig. 14*b*) the maximum in the valence band and the minimum in the conduction band occur at the same momentum ($p = 0$). Thus, an electron making a transition from the valence band to the conduction band can do so without a change in momentum.

Gallium arsenide is called a *direct semiconductor* because it does not require a change in momentum for an electron transition from the valence band to the conduction band. Silicon is called an *indirect semiconductor* because a change of momentum is required in a transition. This difference between direct and indirect band structures is very important for light-emitting diodes and semiconductor lasers. These devices require direct semiconductors to generate efficiently photons (see Chapters 9 and 10).

We can obtain the effective mass from Fig. 14 using Eq. 5. For example, for gallium arsenide with a very narrow conduction–band parabola, the electron effective mass is 0.063 m_0, while for silicon, with a wider conduction–band parabola, the electron effective mass is 0.19 m_0.

1.4.3   Conduction in Metals, Semiconductors, and Insulators

The enormous variation in electrical conductivity of metals, semiconductors, and insulators shown in Fig. 1 may be explained qualitatively in terms of their energy bands. Figure 15 shows the energy band diagrams of three classes of solids—metals, semiconductors, and insulators.
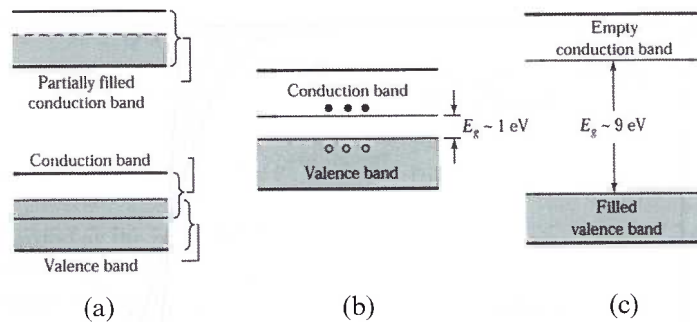
Fig. 15 Schematic energy band representations of (*a*) a conductor with two possibilities (either the partially filled conduction band shown at the upper portion or the overlapping bands shown at the lower portion), (*b*) a semiconductor, and (*c*) an insulator.

Metals

The characteristics of a metal (also called a conductor) include a very low value of resistivity and a conduction band that either is partially filled (as in Cu) or overlaps the valence band (as in Zn or Pb) so that there is no bandgap, as shown in Fig. 15*a*. As a consequence, the uppermost electrons in the partially filled band or electrons at the top of the valence band can move to the next higher available energy level when they gain kinetic energy (e.g., from an applied electric field). Electrons are free to move with only a small applied field in a metal because there are many unoccupied states close to the occupied energy states. Therefore, current conduction can readily occur in conductors.

Insulators

In an insulator such as silicon dioxide (SiO_2), the valence electrons form strong bonds between neighboring atoms. Since these bonds are difficult to break, there are no free electrons to participate in current conduction at or near room temperature. As shown in the energy band diagram (Fig. 15*c*), insulators are characterized by a large bandgap. Note that electrons occupy all energy levels in the valence band and all energy levels in the conduction band are empty. Thermal energy[§] or the energy of an applied electric field is insufficient to raise the uppermost electron in the valence band to the conduction band. Thus, although an insulator has many vacant states in the conduction band that can accept electrons, so few electrons actually occupy conduction band states that the overall contribution to electrical conductivity is very small, resulting in a very high resistivity. Therefore, silicon dioxide is an insulator; it can not conduct current.

Semiconductors

Now, consider a material that has a much smaller energy gap, on the order of 1 eV (Fig. 15*b*). Such materials are called semiconductors. At $T = 0$ K, all electrons are in the valence band, and there are no electrons in the conduction band. Thus, semiconductors are poor conductors at low temperatures. At room temperature and under normal atmospheres, values of E_g are 1.12 eV for Si and 1.42 eV for GaAs. The thermal energy kT at room temperature is a good fraction of E_g, and appreciable numbers of electrons are thermally excited from the valence band to the conduction band. Since there are many empty states in the conduction band, a small applied potential can easily move these electrons, resulting in a moderate current.

▶ 1.5 INTRINSIC CARRIER CONCENTRATION

We now derive the carrier concentration in the thermal equilibrium condition, that is, the steady-state condition at a given temperature without any external excitations such as light, pressure, or an electric field. At a given temperature, continuous thermal agitation results in the excitation of electrons from the valence band to the conduction band and leaves an equal number of holes in the valence band. An *intrinsic semiconductor* is one that contains relatively small amounts of impurities compared with the thermally generated electrons and holes.

---

§ The thermal energy is of the order of kT. At room temperature, kT is 0.026 eV, which is much smaller than the bandgap of an insulator.
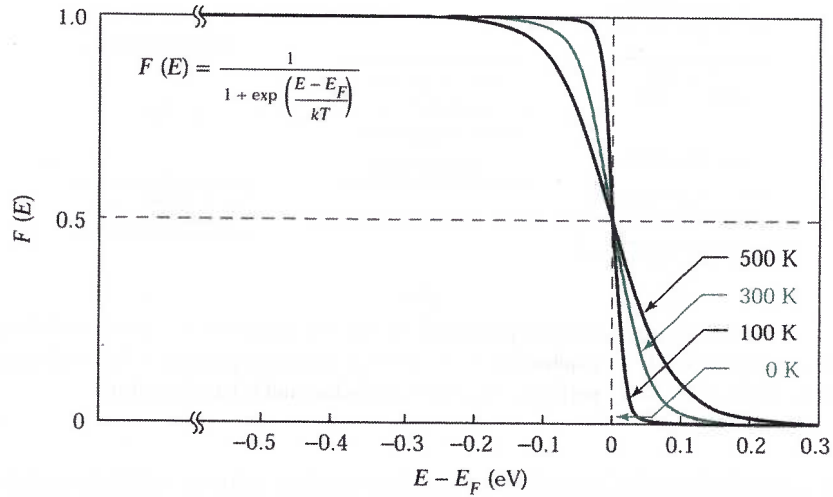
Fig. 16 **Fermi distribution function $F(E)$ versus $(E - E_F)$ for various temperatures.**

To obtain the electron density (i.e., the number of electrons per unit volume) in an intrinsic semiconductor, we first evaluate the electron density in an incremental energy range dE. This density $n(E)$ is given by the product of the density of states $N(E)$, that is, the density of allowed energy states (including electron spin) per energy range per unit volume,[§] and by the probability of occupying that energy range $F(E)$. Thus, the electron density in the conduction band is given by integrating $N(E) F(E) dE$ from the bottom of the conduction band (E_c initially taken to be $E = 0$ for simplicity) to the top of the conduction band E_{top}:

$$n = \int_0^{E_{top}} n(E)\,dE = \int_0^{E_{top}} N(E) F(E)\,dE, \tag{6}$$

where n is in cm^{-3}, and $N(E)$ is in $(cm^3\text{–}eV)^{-1}$.

The probability that an electron occupies an electronic state with energy E is given by the Fermi–Dirac distribution function, which is also called the Fermi distribution function

$$F(E) = \frac{1}{1 + e^{(E-E_F)/kT}}, \tag{7}$$

where k is the Boltzmann constant, T is the absolute temperature in degrees Kelvin, and E_F is the energy of the Fermi level. The Fermi level is the energy at which the probability of occupation by an electron is exactly one-half. The Fermi distribution is illustrated in Fig. 16 for different temperatures. Note that $F(E)$ is symmetrical around the Fermi level E_F.

For energies that are $3kT$ above or below the Fermi energy, the exponential term in Eq. 7 becomes larger than 20 or smaller than 0.05, respectively. The Fermi distribution function can thus be approximated by simpler expressions:

[§] The density of states $N(E)$ is derived in Appendix H.