جلسه ۴

CHAPTER 1

Energy Bands and Carrier Concentration in Thermal Equilibrium

1.2.2 The Diamond Structure

- The element semiconductors, silicon and germanium, have the diamond lattice structure shown in Fig. 4a.

- This structure also belongs to the fcc crystal family and can be seen as two interpenetrating(درهم فرو رفتن) fcc sublattices with one sublattice displaced from the other by one-quarter of the distance along the body diagonal of the cube.

- Although chemically identical, the two sets of atoms belonging to the two sublattices are different in terms of the crystal structure.

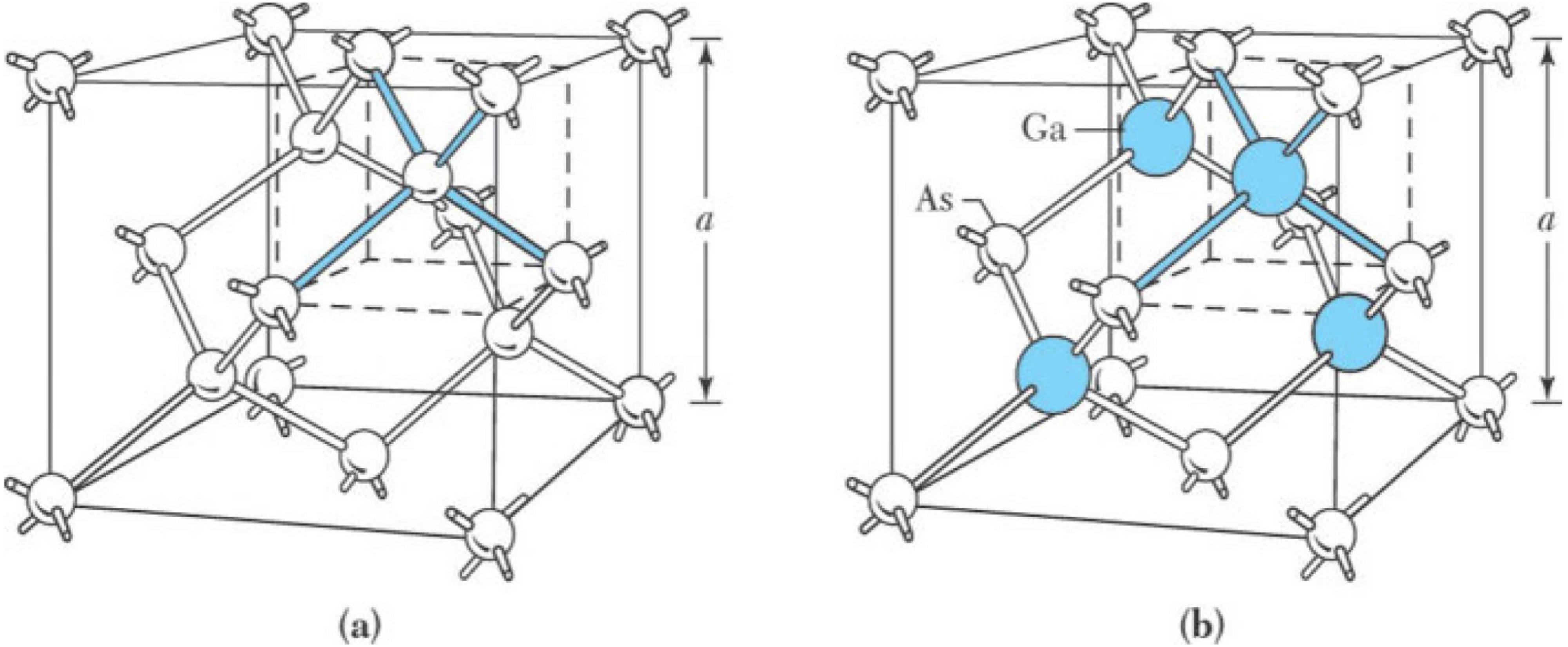


Fig. 4 (*a*) Diamond lattice. (*b*) Zincblende lattice.

- A unit cell of a diamond lattice consists of a tetrahedron(چهار وجهی) in which each atom is surrounded by four equidistant nearest neighbors that lie at the corners (the spheres connected by darkened bars in Fig. 4a).

- Most of the III-V compound semiconductors (e.g., GaAs) have a zincblende lattice, shown in Fig. 4b, which is identical to a diamond lattice except that one fcc sublattice has Column III atoms (Ga) and the other has Column V atoms (As).

- Appendix F gives a summary of the lattice constants and other properties of important element and binary compound semiconductors.

EXAMPLE

At 300 K the lattice constant for silicon is 5.43 Å. Calculate the number of silicon atoms per cubic centimeter and the density (mass of a unit volume)of silicon at room temperature.

1.2.3 Crystal Planes and Miller Indices

- In Fig. 3b we note that there are four atoms in the ABCD plane and five atoms in the ACEF plane (four atoms from the corners and one from the center) and that the atomic spacing is different in the two planes. Therefore, the crystal properties along different planes are different, and the electrical and other device characteristics can be dependent on the crystal orientation.
- A convenient method of defining the various planes in a crystal is to use Miller indices.
- These indices are obtained using the following steps:

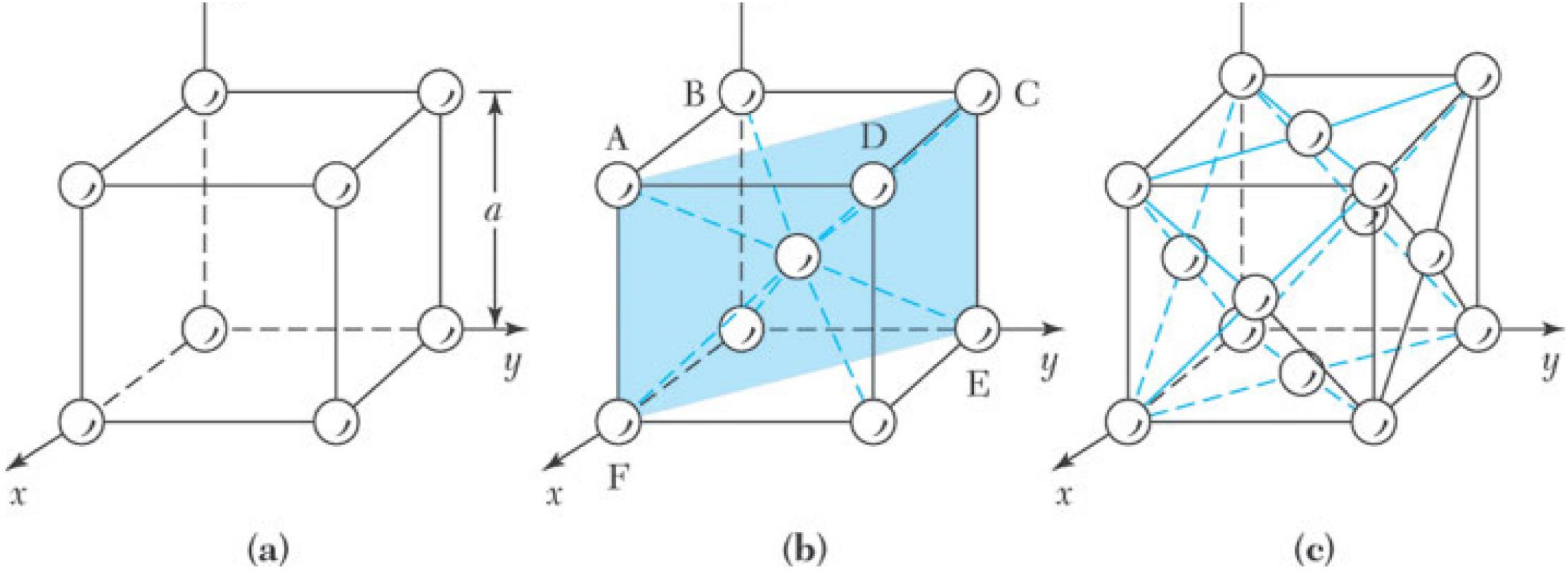


Fig. 3 **Three cubic-crystal unit cells. (*a*) Simple cubic. (*b*) Body-centered cubic. (*c*) Face-centered cubic.**

1. Find the intercepts(تقاطع) of the plane on the three Cartesian coordinates in terms of the lattice constant.

2. Take the reciprocals of these numbers and reduce them to the smallest three integers having the same ratio.

3. Enclose the result in parentheses (hkl) as the Miller indices for a single plane.

EXAMPLE

In Fig. 5, find the Miller indices for a specified plane.

SOLUTION

As shown in Fig. 5, the plane has intercepts at a, 3a, and 2a along the three coordinates. Taking the reciprocals of these intercepts, we get 1, 1⁄3, and 1⁄2. The smallest three integers having the same ratio are 6, 2, and 3 (obtained by multiplying each fraction by 6). Thus, the plane is referred to as a (623)-plane.

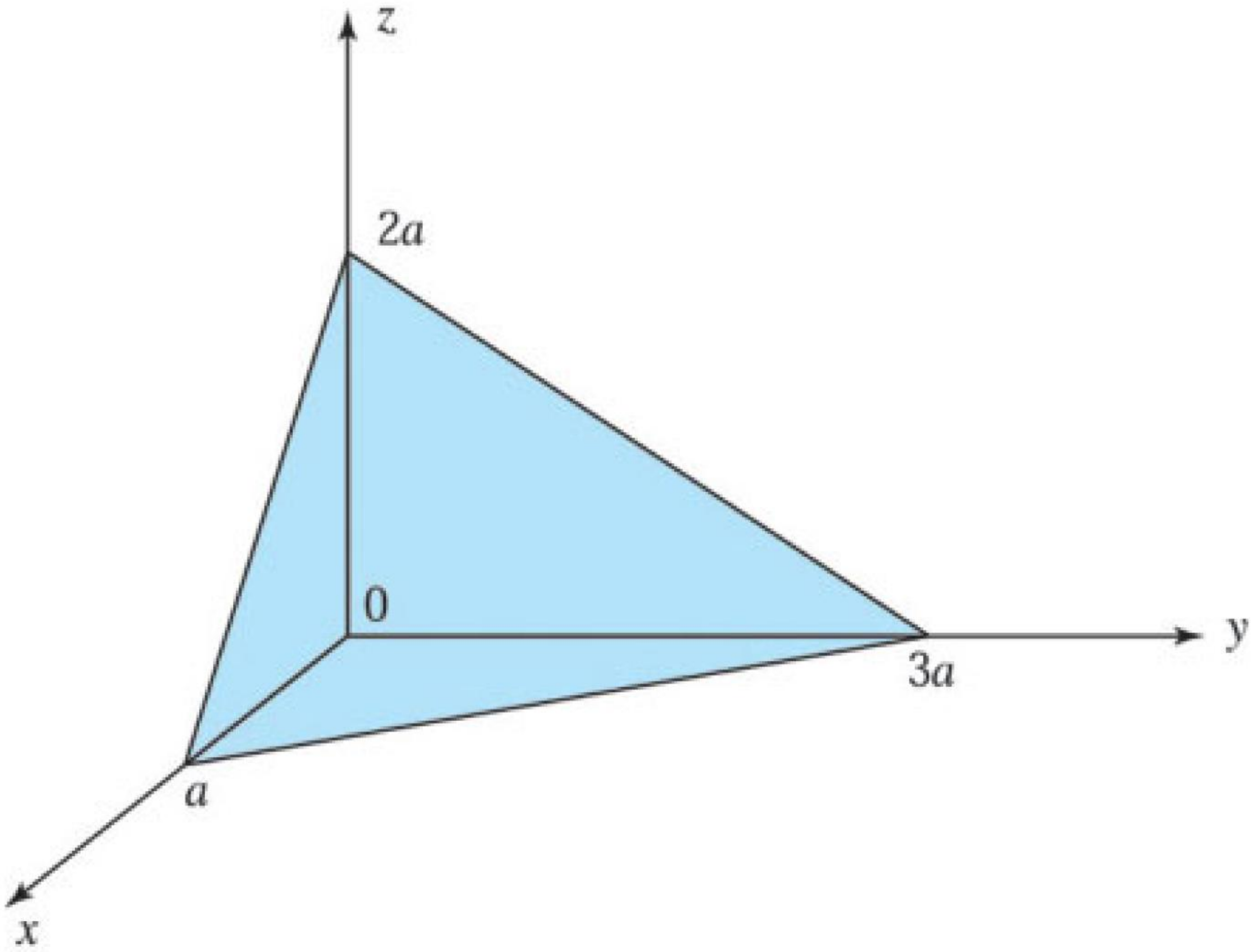
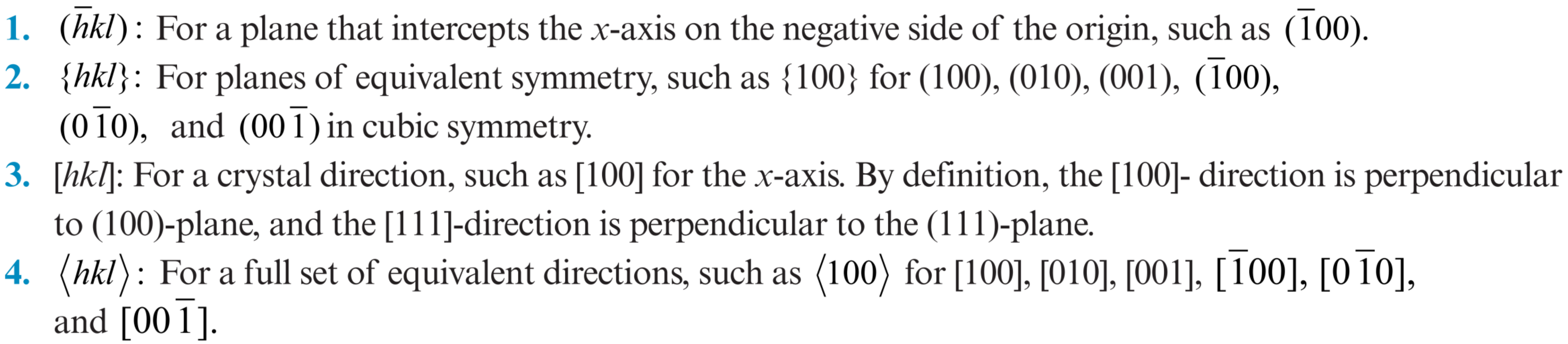


Fig. 5 **A (623)-crystal plane.**

Figure 6 shows the Miller indices of important planes in a cubic crystal.§ Some other conventions are the following:



1. $(\bar{h}kl)$: For a plane that intercepts the x-axis on the negative side of the origin, such as $(\bar{1}00)$.
2. $\{hkl\}$: For planes of equivalent symmetry, such as $\{100\}$ for (100), (010), (001), $(\bar{1}00)$, $(0\bar{1}0)$, and $(00\bar{1})$ in cubic symmetry.
3. $[hkl]$: For a crystal direction, such as [100] for the x-axis. By definition, the [100]- direction is perpendicular to (100)-plane, and the [111]-direction is perpendicular to the (111)-plane.
4. $\langle hkl \rangle$: For a full set of equivalent directions, such as $\langle 100 \rangle$ for [100], [010], [001], $[\bar{1}00]$, $[0\bar{1}0]$, and $[00\bar{1}]$.

§ In Chapter 5, we show that the <100> orientation is preferred for silicon metal-oxide- semiconductor field-effect transistors (MOSFETs).

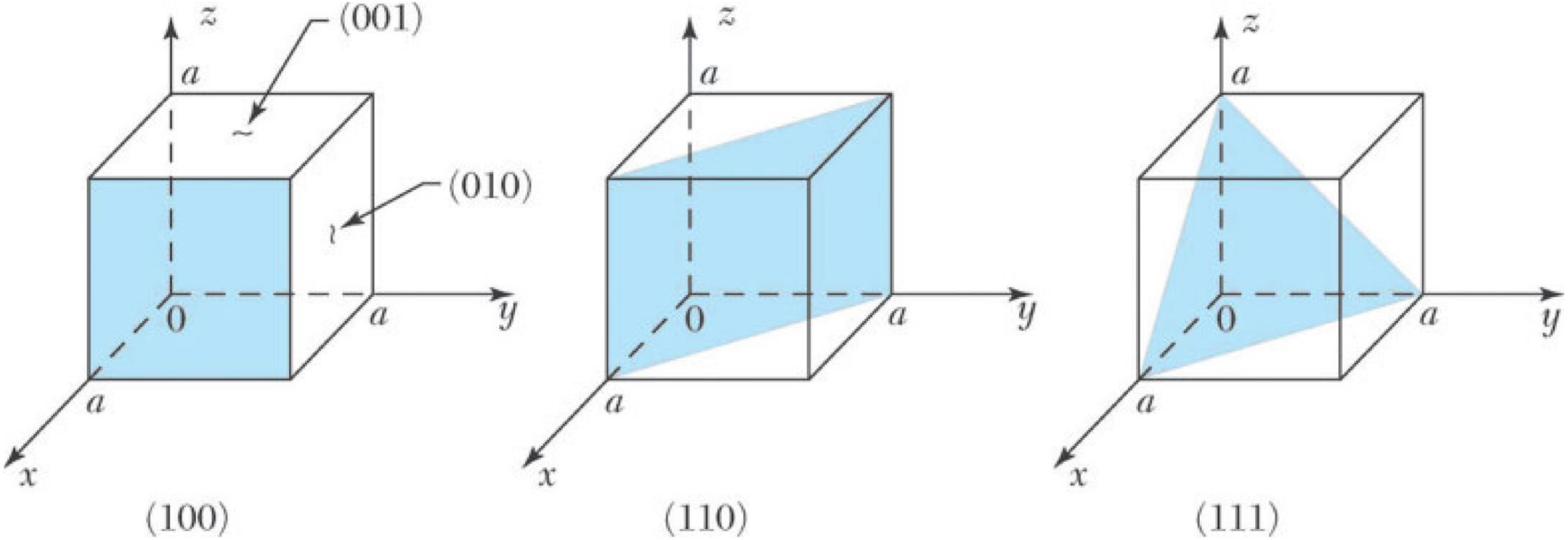


Fig. 6 **Miller indices of some important planes in a cubic crystal.**

1.3 VALENCE BONDS

- As discussed in Section 1.2, each atom in a diamond lattice is surrounded by four nearest neighbors.
- Figure 7a shows the tetrahedron bonds of a diamond lattice.
- A simplified two-dimensional bonding diagram for the tetrahedron is shown in Fig. 7b.
- Each atom has four electrons in the outer orbit, and each atom shares these valence electrons with its four neighbors. This sharing of electrons is known as covalent bonding; each electron pair constitutes a covalent bond.

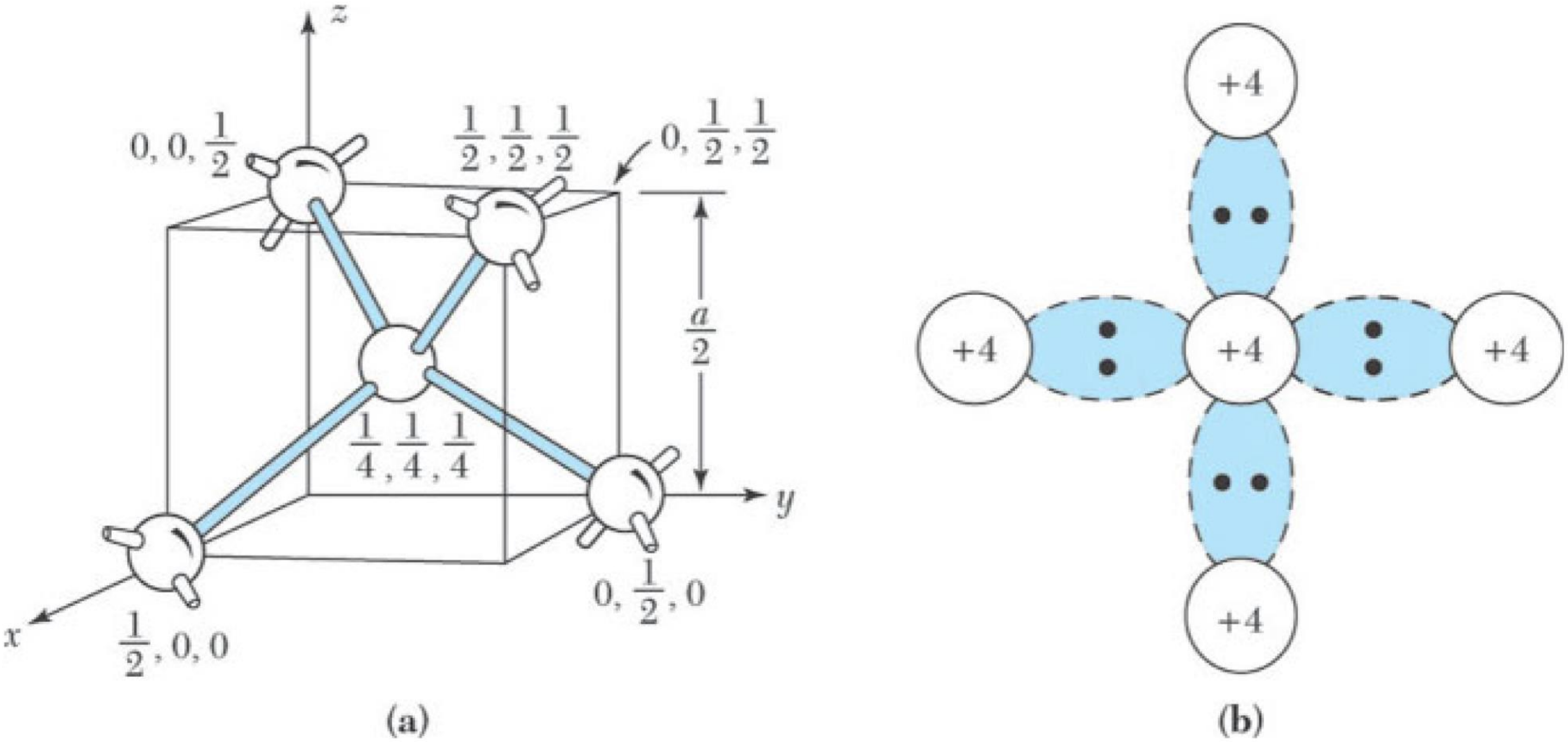


Fig. 7 **(*a*) A tetrahedron bond. (*b*) Schematic two-dimensional representation of a tetrahedron bond.**

- Covalent bonding occurs between atoms of the same element or between atoms of different elements that have similar outer-shell electron configurations.

- Each electron spends an equal amount of time with each nucleus. However, both electrons spend most of their time between the two nuclei.

- The force of attraction for the electrons by both nuclei holds the two atoms together.

- Gallium arsenide crystallizes in a zincblende lattice, which also has tetrahedron bonds.

- The major bonding force in GaAs is also due to the covalent bond.

- However, gallium arsenide has a small ionic contribution that is an electrostatic attractive force between each Ga^+ ion and its four neighboring As^- ions, or between each As^- ion and its four neighboring Ga^+ ions.

- Electronically, this means that the paired bonding electrons spend slightly more time in the As atom than in the Ga atom.

- At low temperatures, the electrons are bound in their respective tetrahedron lattice; consequently, they are not available for conduction.

- At higher temperatures, thermal vibrations may break the covalent bonds (ionize one electron from the bond).

- When a bond is broken, a free electron results and can participate in current conduction.

- Figure 8a shows the situation when a valence electron in silicon becomes a free electron.

- An electron deficiency(کمبود) is left in the covalent bond. This deficiency may be filled by one of the neighboring electrons, which results in a shift of the deficiency location, as from location A to location B in Fig. 8b.

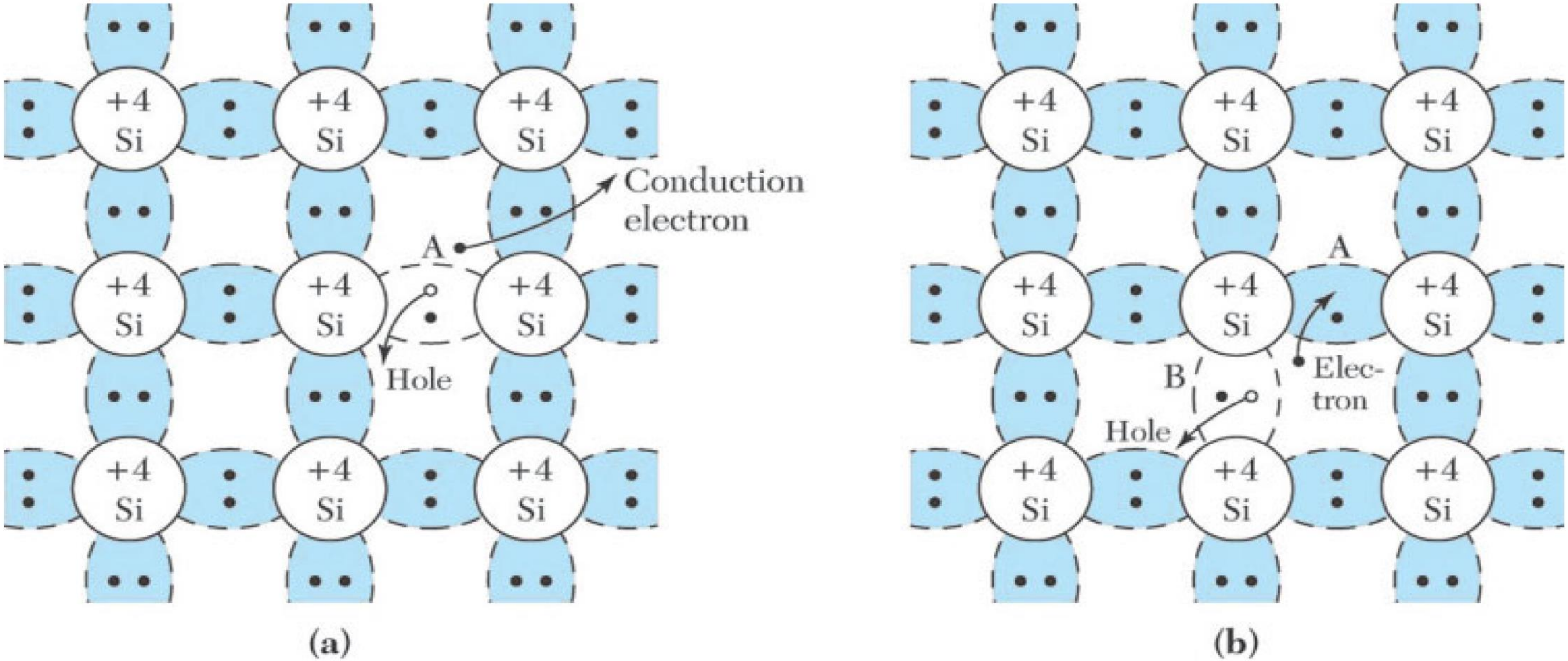


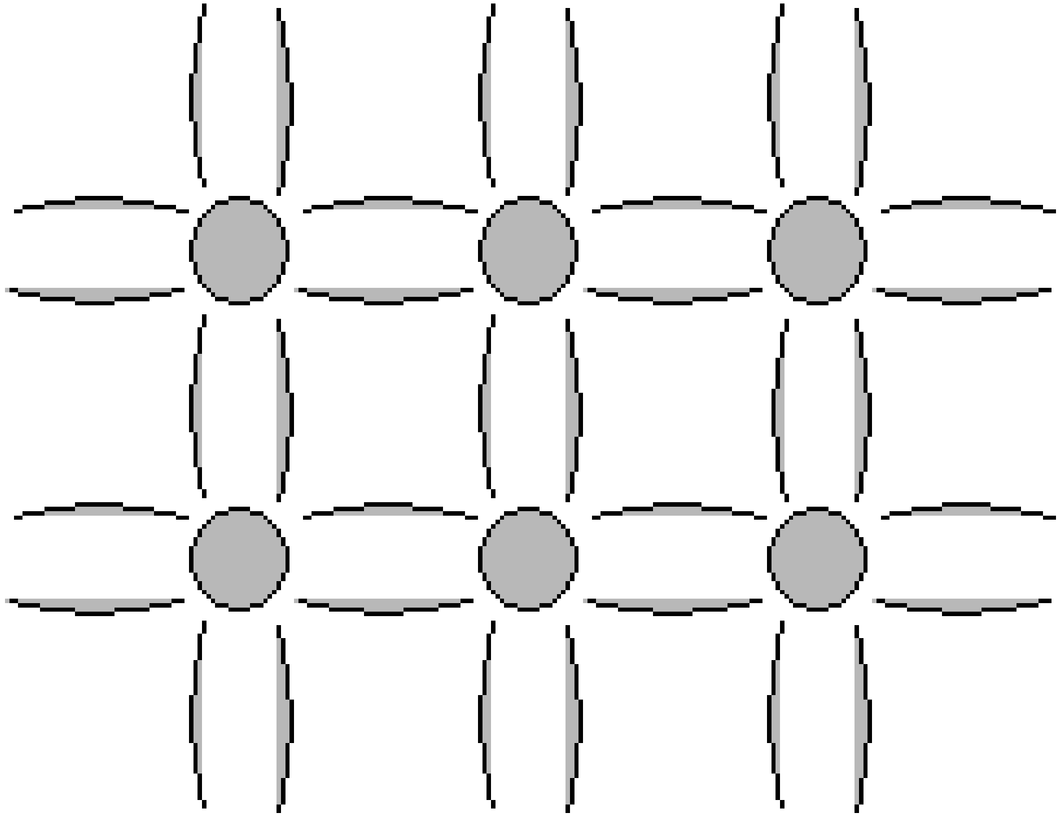
Fig. 8 **The basic bond representation of intrinsic silicon. (*a*) A broken bond at position A, resulting in a conduction electron and a *hole*. (*b*) A broken bond at position B.**

- We may, therefore, consider this deficiency as a particle similar to an electron.

- This fictitious (ساختگی، خیالی)particle is called a hole. It carries a positive charge and moves, under the influence of an applied electric field, in the direction opposite to that of an electron.

- Therefore, both the electron and the hole contribute to the total electric current.

- The concept of a hole is analogous to that of a bubble in a liquid: although it is actually the liquid that moves, it is much easier to talk about the motion of the bubble in the opposite direction.



Animation showing formation of "free" electrons and holes when an electron can escape its bond.

1.4 ENERGY BANDS

1.4.1 Energy Levels of Isolated Atoms

- For an isolated atom, the electrons can have discrete energy levels. For example, the energy levels for an isolated hydrogen atom are given by the Bohr model:

$$E_H = -m_0 q^4 / 8\varepsilon_0{}^2 h^2 n^2 = -13.6 / n^2 \ \ \text{eV}, \tag{2}$$

where m_0 is the free-electron mass, q is the electronic charge, ε_0 is the free-space permittivity, h is the Planck constant, and n is a positive integer called the principal quantum number.

The quantity eV (electron volt) is an energy unit corresponding to the energy gained by an electron when its potential is increased by one volt. It is equal to the product of q (1.6×10^{-19} coulomb) and one volt, or 1.6×10^{-19} J. The discrete energies are –13.6 eV for the ground- state energy level ($n = 1$), – 3.4 eV for the first excited-state energy level ($n = 2$), and so on. Detailed studies reveal that for higher principle quantum numbers ($n \geq 2$), energy levels are split according to their angular momentum quantum number ($\ell = 0, 1, 2, \ldots, n - 1$).

- We now consider two identical atoms.

- When they are far apart, the allowed energy levels for a given principal quantum number (e.g., n = 1) consist of one doubly degenerate level; that is, both atoms have exactly the same energy.

- When they are brought closer, the doubly degenerate energy levels will spilt into two levels by the interaction between the atoms.

- The split occurs due to the Pauli exclusion principle, which states that no more than two electrons in a given system can reside(واقع شدن) in the same energy state at the same time.

- As N isolated atoms are brought together to form a solid, the orbits of the outer electrons of different atoms overlap and interact with each other.
- This interaction, including those forces of attraction and repulsion between atoms, causes a shift in the energy levels, as in the case of two interacting atoms. However, instead of two levels, N separate but closely spaced levels are formed.
- When N is large, the result is an essentially continuous band of energy. This band of N levels can extend over a few eV at the inter-atomic distance of the crystal.

- The electrons can no longer be treated as belonging to their parent atoms. They belong to the crystal as a whole.

- Figure 9 shows the effect, where the parameter a represents the equilibrium inter-atomic distance of the crystal.

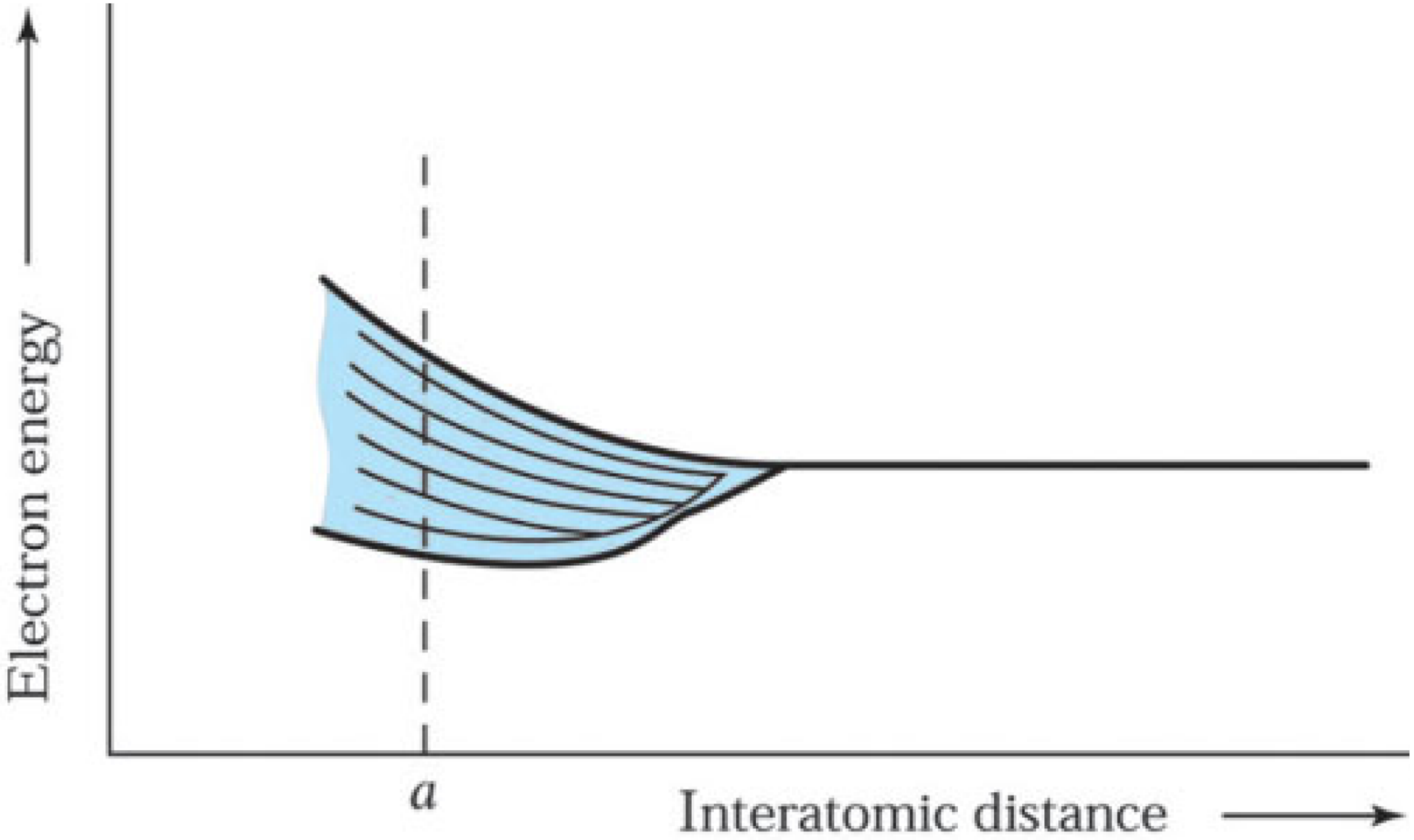


Fig. 9 **The splitting of a degenerate state into a band of allowed energies.**

- The actual band splitting in a semiconductor is much more complicated.

- Figure 10 shows an isolated silicon atom that has 14 electrons.

- Of the 14 electrons, 10 occupy deep-lying energy levels whose orbital radius is much smaller than the interatomic separation in the crystal.

- The four remaining valence electrons are relatively weakly bound and can be involved in chemical reactions.

- Therefore, we only need to consider the outer shell (the n =3 level) for the valence electrons, since the two inner shells are completely full and tightly bound to the nucleus.

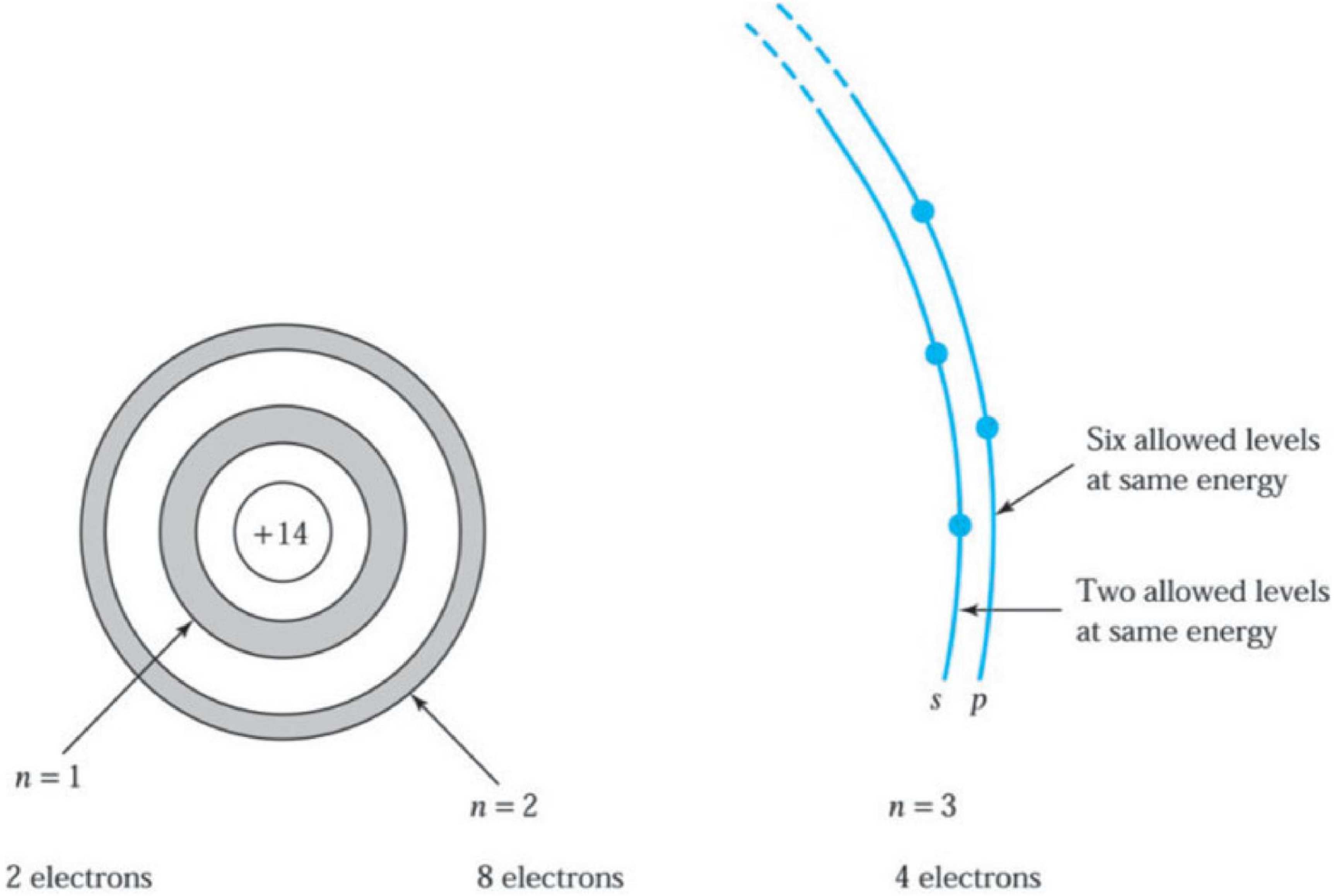


Fig. 10 **Schematic representation of an isolated silicon atom.**

- The 3s subshell (i.e., for n = 3 and ℓ = 0) has two allowed quantum states per atom. This subshell will contain two valence electrons at T = 0 K.

- The 3p subshell (i.e., n = 3 and ℓ = 1) has six allowed quantum states per atom. This subshell will contain the remaining two valence electrons of an individual silicon atom.

- Figure 11 is a schematic diagram of the formation of a silicon crystal from N isolated silicon atoms.

- As the interatomic distance decreases, the 3s and 3p subshell of the N silicon atoms will interact and overlap to form bands.

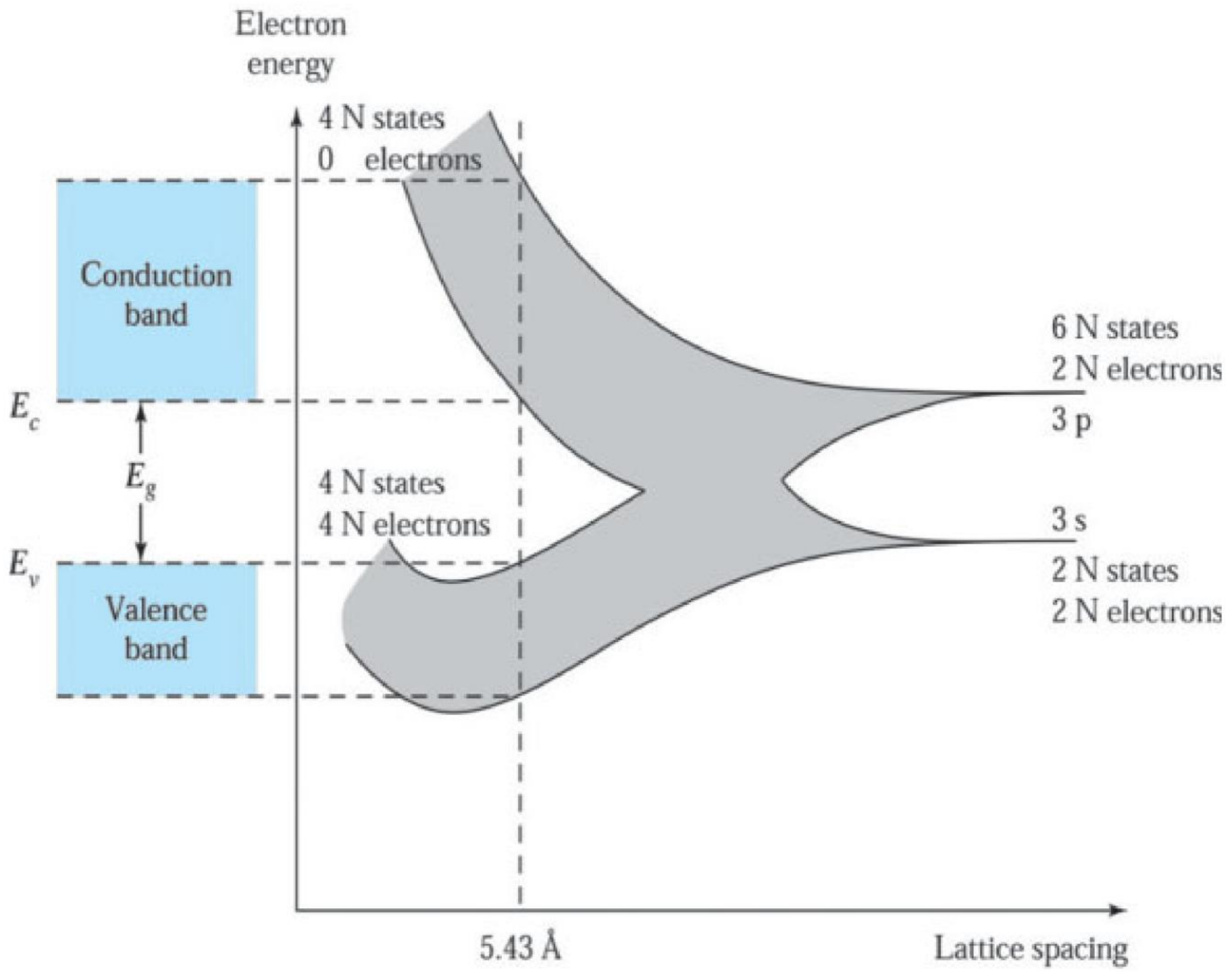


Fig. 11 **Formation of energy bands as a diamond lattice crystal is formed by bringing isolated silicon atoms together.**

- As the 3s and 3p bands grow, they merge(ترکیب شدن) into a single band containing 8 N states.
- At the equilibrium interatomic distance determined by the condition of minimum total energy, the bands will again split, with 4 N states in the lower band and 4 N states in the upper band.
- At a temperature of absolute zero, electrons occupy the lowest energy states, so that all states in the lower band (the valence band) will be full and all states in the upper band (the conduction band) will be empty.

- The bottom of the conduction band is called E_C and the top of the valence band is called E_V.

- The bandgap energy E_g between the bottom of the conduction band and the top of the valence band ($E_C - E_V$) is the width of the forbidden energy gap, as shown at the far left of Fig. 11.

- Physically, E_g is the energy required to break a bond in the semiconductor to free an electron to the conduction band and leave a hole in the valence band.

1.4.2 The Energy-Momentum Diagram

- The energy E of a free electron is given by

$$E = \frac{p^2}{2m_0}, \qquad (3)$$

where p is the momentum and m_0 is the free-electron mass.

- If we plot E vs. p, we obtain a parabola as shown in Fig. 12.

- In a semiconductor crystal, an electron in the conduction band is similar to a free electron in being relatively free to move about in the crystal. However, because of the periodic potential of the nuclei, Eq. 3 can no longer be valid.

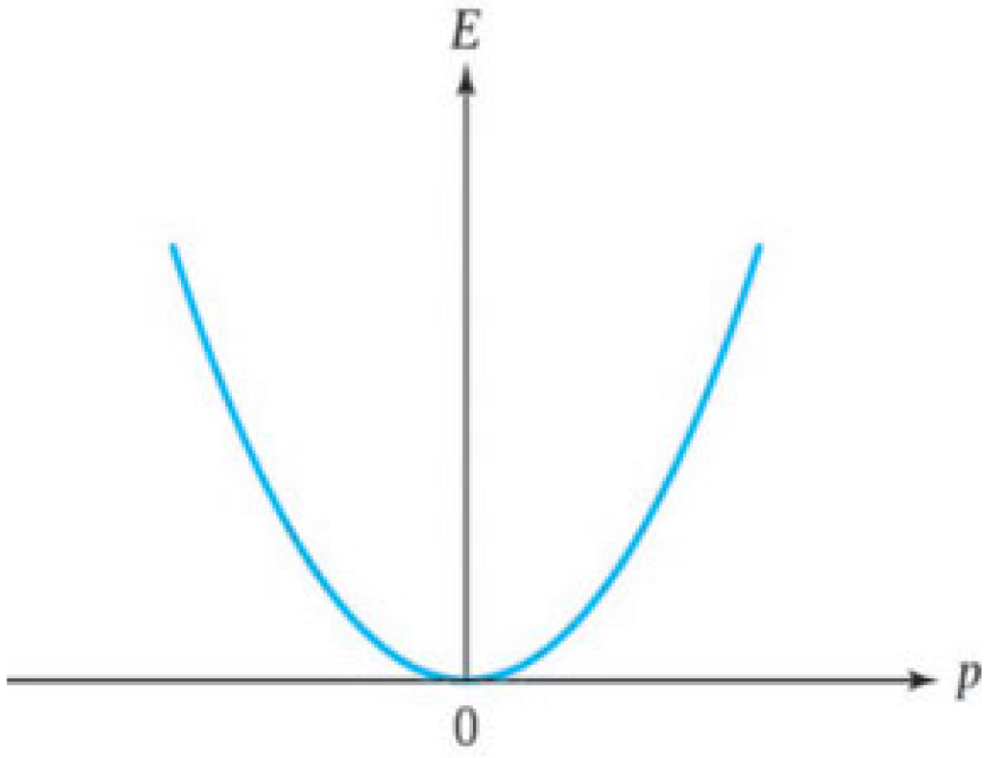


Fig. 12 The parabolic energy (E) vs. momentum (p) curve for a free electron.

- However, it turns out that we can still use Eq. 3 if we replace the free-electron mass in Eq. 3 by an effective mass m_n (the subscript n refers to the negative charge on an electron), that is,

$$E = \frac{p^2}{2m_n}. \qquad (4)$$

- The electron effective mass depends on the properties of the semiconductor.

- If we have an energy-momentum relationship described by Eq. 4, we can obtain the effective mass from the second derivative of E with respect to p:

$$m_n \equiv \left(\frac{d^2E}{dp^2} \right)^{-1}. \qquad (5)$$

- Therefore, the narrower the parabola, corresponding to a larger second derivative, the smaller the effective mass. ↓

- A similar expression can be written for holes (with effective mass m_p where the subscript p refers to the positive charge on a hole).

- The effective-mass concept is very useful because it enables us to treat electrons and holes essentially as classical charged particles.

- Figure 13 shows a simplified energy-momentum relationship of a special semiconductor with an electron effective mass of $m_n = 0.25m_0$ in the conduction band (the upper parabola) and a hole effective mass of $m_p = m_0$ in the valence band (the lower parabola).

- Note that the electron energy is measured upward and the hole energy is measured downward.

- The spacing at p = 0 between these two parabolas is the bandgap Eg, shown previously in Fig. 11.

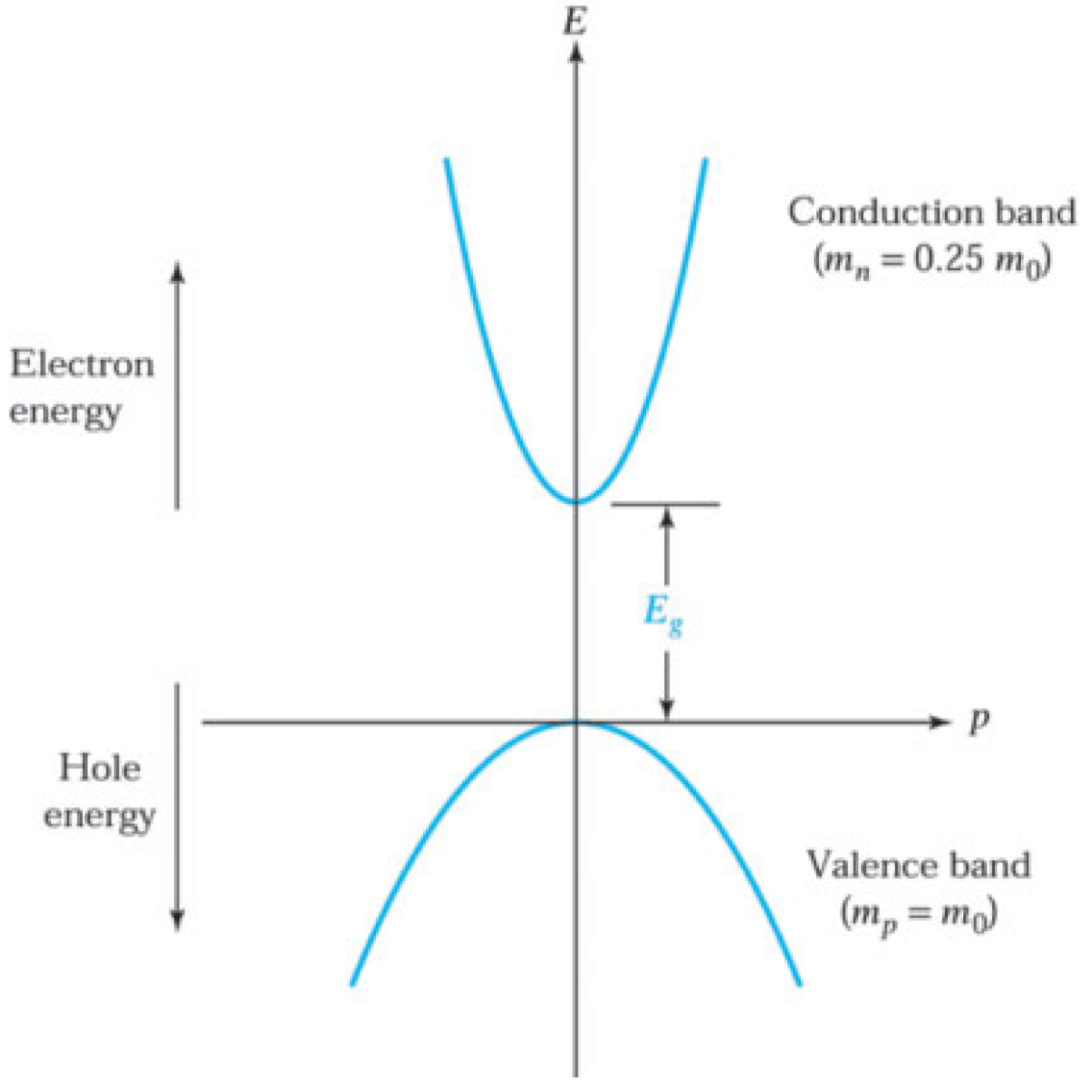


Fig. 13 **A schematic energy-momentum diagram for a special semiconductor with $m_n = 0.25\ m_0$ and $m_p = m_0$.**

- The actual energy-momentum relationships (also called energy-band diagram) for silicon and gallium arsenide are much more complex.

- Visualized in three dimensions, the relationship between E and p is a complex surface.

- They are shown in Fig. 14 only for two crystal directions.

- Since the periodicity of most lattice is different in various directions, the energy-momentum diagram is also different for different directions.

- In the case of the diamond or zincblende lattice, the maximum in the valence band and minimum in the conduction band occur at $p = 0$ or along one of these two directions.

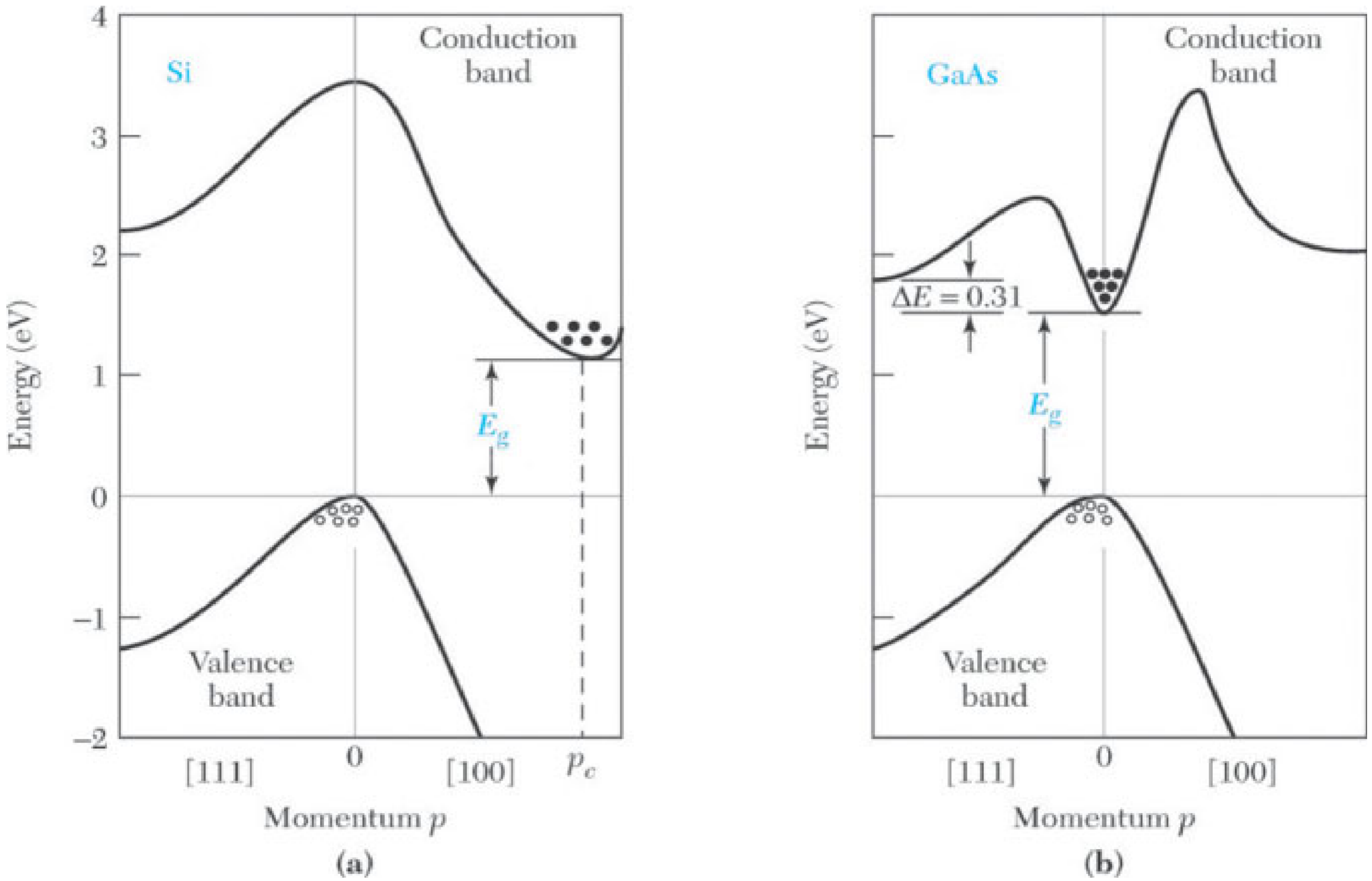


Fig. 14 **Energy band structures of (*a*) Si and (*b*) GaAs. Circles (o) indicate holes in the valence bands and dots (•) indicate electrons in the conduction bands.**

- If the minimum of the conduction band occurs at $p = 0$, this means the effective mass of the electrons in every direction in the crystal is the same. It also indicates that the electron motion is independent of crystal direction.

- If the minimum of the conduction band occurs at $p \neq 0$, this means that the electron behavior in every direction is not the same in the crystal.

- In general, the minimum of conduction band of polar (with partly ionic binding) semiconductors tend to be at $p = 0$, which is related to the lattice structure and the fraction of ionicity in the bond.

- 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. 14a) 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 ($\geq$ Eg) but also some momentum change ($\geq p_c$) is required.

- For gallium arsenide (Fig. 14b) 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 $0.063m_0$, while for silicon, with a wider conduction–band parabola, the electron effective mass is $0.19m_0$.